

小学校国語科の教科書における説明文の品詞の分析

Word class analysis of explanatory texts in primary school Japanese
language textbooks

中井 悠加 藤木 大介 井上 弥
(保育教育学科) (広島大学大学院人間社会科学研究科)

キーワード：説明文教材、品詞、小学校国語科、分類木分析

1. はじめに

国語科における説明文は、「ある物事の成り立ちや仕組みを、その物事に関して未知な読み手に対してその理解を得るように一定のものの見方・考え方から説き明かして伝える文章（大内 2011, p.292）」とされ、国語科の「読むこと」として重要な教材である。近年の国語教育学における説明文の教材研究は、論理・論証に焦点を当てて展開されており（難波・三原市立木原小学校 2009；光野 2005；間瀬 2017；古賀 2020）、そのような読む力は、ブルームの教育目標のタクソノミーにおける高次の能力に相当する「高次読解力」と位置づけられる（間瀬 2017；石井 2002）。説明文教材の表現や構造に着目した研究としては、例えば井上（1986）は、学校図書（昭和 54 年度版）中の説明文 25 教材を対象に、説明文の語彙構造を検討している。本文の長さを反映する延べ語数と語彙の数を反映する異なり語数の学年別変化から両者が比例していること、また変化率から 1 年生、2 年生、3 年生、4 年生以上の高学年の 4 段階に分けられることを見出している。また船所（2014）は、5 社の教科書の説明文を基に、使用される題材や比較・因果などの論理的思考の種類を中心に内容面からの学年比較をしている。さらに岸ほか（1989）は、教育出版（昭和 58 年度版）を対象に、文章構成の観点から学年ごとの説明文の類型を行った。

このように、国語科の教科書を分析した研究では語彙、内容、文章の構造等の分析が行われてきた。しかしながら、それらの表現形式・表現方法の特徴が冒頭で挙げた「高次読解力」の基礎を養う役割としてどう位置付くのかは検証されていない。例えば、先般児童生徒の読解力の低下を指摘したリーディングスキルテストでは、係り受けの分析（依存関係分析：dependency analysis）や、代名詞や代名詞の省略されたゼロ代名詞が指示する対象を同定する照応解決（anaphora resolution）や、句の境界を示す格助詞「に」の並列とこれに対応する 1 つの述語との依存関係などを要する短文の理解度を測っているが（ARAI et al. 2017；国立情報学研究所 2016；新井 2018）、それらの表現がど

表1 教材のタイトル

学年	タイトル			
1	うみのかくれんぼ (11文, 317文字)	くちばし (16文, 362文字)	じどう車くらべ (12文, 395文字)	どうぶつの赤ちゃん (19文, 704文字)
2	おにごっこ (28文, 1132文字)	しかけカードのつくり方 (34文, 810文字)	たんぼぼのちえ (18文, 711文字)	どうぶつ園のじゅうい (41文, 1384文字)
3	ありの行列 (32文, 1048文字)	こまを楽しむ (34文, 1163文字)	すがたをかえる大豆 (39文, 1255文字)	言葉で遊ぼう (21文, 791文字)
4	アップとルーズ (35文, 1146文字)	ウナギのなぞ (61文, 2517文字)	大きな力を出す (19文, 637文字)	動いて、考えて、また動く (39文, 1561文字)
5	見立てる (19文, 666文字)	生き物は円柱形 (58文, 1603文字)	想像力のスイッチ (56文, 2344文字)	天気を予想する (45文, 2030文字)
6	『鳥獣戯画』を読む (67文, 2211文字)	時計の時間と心の時間 (53文, 2360文字)	自然に学ぶ暮らし (55文, 2242文字)	笑うから楽しい (16文, 700文字)

の程度国語教科書に用いられているかはこれまでの国語教科書研究では検討の対象となっていない。国語科の教科書を通して触れる表現について知る手がかりの1つにこの格助詞などの品詞がある。計量文体学と呼ばれる領域では、文の長さや品詞の使用率、特定の単語の使用率、語彙の豊富さなどに基づいて、文学作品の作者の同定などが行われる（例えば、村上 2000 参照）。また、文章の種類毎に比較したものもあり、石田ほか（2004）は、論文は新聞、日記に比べて助詞、動詞、助動詞の割合が小さいことを示している。

そこで本研究では、小学校国語教科書の説明文における品詞の使用率に着目し、学年毎に特徴があるか明らかにする。また、これに基づき、学年毎の特徴から児童がどのような修辞表現に多く触れているか調べ、教科書を通して係り受けの分析や照応解決に必要な能力が育まれるように構成されているか検討する。例えば、係り受けについて段階的に学べるよう、学年が上がるとともに節境界を示すために句読法に変化が見られたり、「～と」「～という」といった助詞の使用頻度が増えたりする可能性がある。また、照応解決については学年とともに代名詞を用いた表現が増えていく可能性も考えられる。さらに、説明文中の品詞の使用率の変化を捉えることで、児童が段階的に学んでいる表現について明らかにする。

2. 方法

対象教科書：1年生から6年生までの小学校国語科の教科書（光村図書 平成28年度版）を対象とした¹⁾。

説明文教材：各学年の教科書から、表1に示したような説明文教材をすべて選択した。いずれの学年も4教材であった。教科書のテキストの電子化は全て執筆者の手入力により行った。その際、教科書教材としての特徴を明らかにする目的から低学年に主に見受けられる空白も全て1文字分として含めている。1年生教材では文中の空白および句読点のない改行箇所を1文字分の空白と判

3. 結果および考察

1) 分析の結果

分析は R 3.4.2 (RCORE TEAM 2017) を用いて行った。

まず、説明文教材を、引用括弧内の句点を除く句点「。」で区切られたものを一文として電子化した。次に、mecab-ipadic-NEologd(佐藤・橋本・奥村 2017) を辞書として、RMeCab パッケージ (ISHIDA 2017) から MeCab (KUDO et al. 2004) を用いて形態素に分解した。学年別に、教材数、文の総数、1教材あたりの文の数、1文当たりの文字数および各品詞・品詞細分類(以下、品詞・細分類の形式で表記)の出現頻度とその学年差を示したものが表 2 である。例えば、1年生教材の「おもいにもつをのせるトラックには、タイヤがたくさんついています。」(1年生・じどう車くらべ)と6年生教材の「このような厳しい気候の中でも、巣の中の温度は、ほぼ三十度に保たれています。」(6年生・自然に学ぶ暮らし)は同じ37文字とカウントしている。

この表 2 の出現頻度を基に、1文中に必ずある品詞や全くない品詞は分散がなくなり(標準偏差が 0.000 になる)検定できないため、すべての学年で教材文による分散がある1文当たりの文字数および品詞について、学年を要因とする1要因の分散分析を行った。その結果を示したものが表 2 の学年差の列である。

いずれも分散の等質性は仮定できないため学年差の検定には Welch の修正を用い、学年差が有意な場合の多重比較は Holm 法を用いた。表 2 のうち、学年による差のあった記号・形容詞・助動詞・副詞・連体詞をまとめて図示したものが図 1、助詞をまとめて図示したものが図 2、動詞をまとめて図示したものが図 3、名詞をまとめて図示したものが図 4 である。

図 1 からわかるように、記号・読点は1年生より6年生が多く、副詞・一般は3年生より4年生が多くなっている。しかし、助動詞、形容詞、連体詞では学年差は有意であるが、学年間の多重比較では有意な違いがみられなかった。次に、助詞をまとめた図 2 からわかるように、格助詞は1年生と比べ2年生から6年生では多く、接続助詞は5年生より6年生が多い。係助詞と並立助詞は2年生よりも6年生が多くなっていた。しかし、終助詞と助詞・連体化では学年差は有意であったが、多重比較では有意な違いがみられなかった。動詞をま

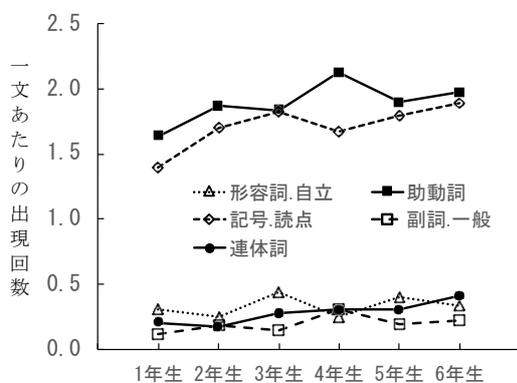


図 1 形容詞・自立、助動詞、記号・読点、副詞・一般、連体詞の学年変化

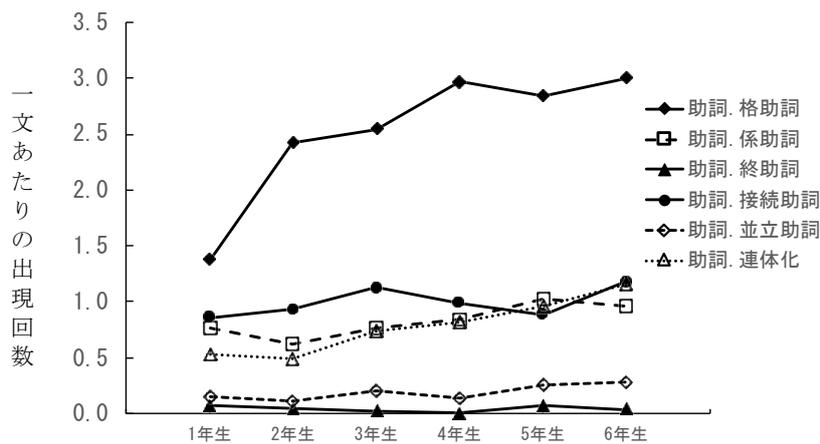


図2 助詞の学年変化

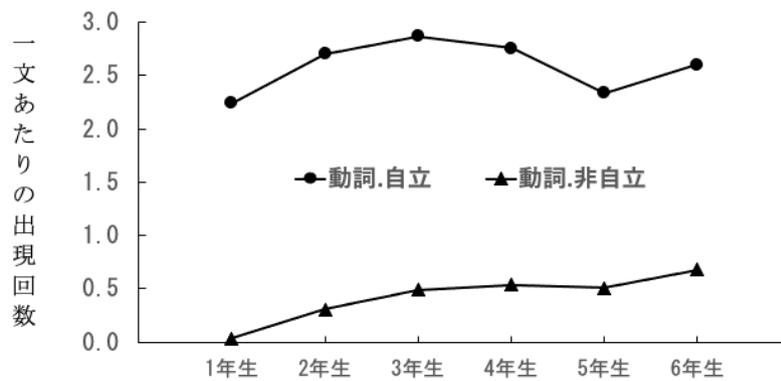


図3 動詞の学年変化

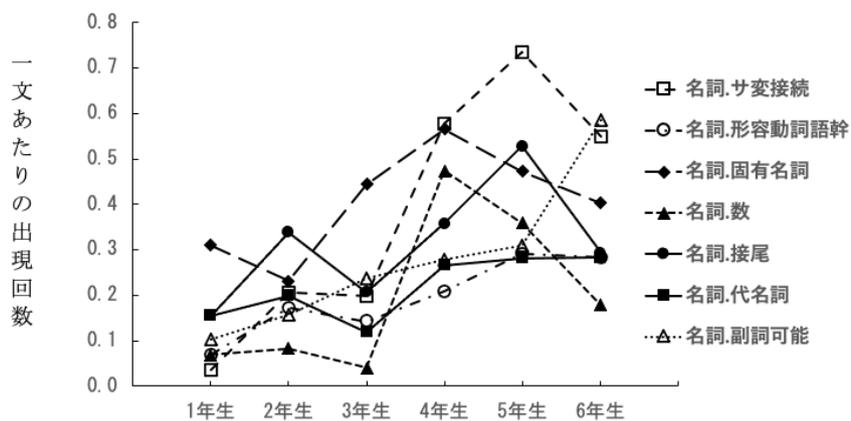


図4 名詞の学年変化

とめた図3からわかるように、動詞.自立は5年生よりも3年生が多く、動詞.非自立は1、2年生より4、6年生が多く、1年生より3、5年生が多くなっていった。さらに、名詞をまとめた図4からもわかるように、名詞.サ変接続は1、2、3年生よりも4、5、6年生が多く、名詞.形容動詞語幹は1、3年生より5、

6年生が多くなっていた。名詞・固有名詞は2年生よりも4、5年生が多く、名詞・代名詞は3年生より5、6年生が多くなっていた。これらの名詞は高学年ほど多いという点で同じ傾向を示していた。しかし、名詞・数は3年生より5年生が多いものの1、2、3、6年生よりも4年生が多くなっていた。また、名詞・代名詞も1、3、6年生よりも5年生が多くなっていた。このように、品詞ごとに使用回数の差が出る学年が異なることから、学年ごとの特徴を考える上では、どの品詞によって学年を識別できるかが鍵となる。

教科書は全般的には学年が上がるにつれて特徴も変化していくようなもので、学年間で連続性を欠く特徴ばかりを有するとは考えにくい。そう考えると、教科書の学年ごとの特徴を示す上では、まず、各学年の教科書の特徴を大きく分ける品詞は何か特定され、段階的に各学年の特徴も特定されていくような分析も必要であると考えた。そこで、各文の学年を基準変数、文字数を除き、抽出した品詞を説明変数として、rpartパッケージ (THERNEAU et al. 2017) による分類木分析を行った。その際、記号・括弧開と記号・括弧閉は必ず対になっているため、記号・括弧閉を説明変数から除き、名詞・アルファベットは5年生の題材で「Aさん」といった表現で用いられた特殊なものなので除いた。交差を足した値と樹形図のわかりやすさを基に、複雑性パラメーター (complexity parameter) 値 0.01255887 で刈り込みを行ったものが、図5の樹形モデルである。

この図5からわかるように、1年生は記号・空白の多さと助詞・格助詞の少なさで、2年生は1年生よりも記号・空白が少なく、助詞・格助詞が多いことで分離できているが、3年生以上になると助動詞や助詞、動詞といった分岐点も増え、より複雑な分岐になっている。

樹形モデルでも、記号・空白の多さによって1、2年生と3年生以上に大きく分かれており、さらに同じ記号・空白の多さで1年生と2年生が分岐している。また、3年生以上では助動詞の少なさが5年生をその他の学年から分け、さらに名詞・サ変接続、名詞・接尾語の多さが5年生を分けている。その後、名詞・副詞可能の多さと助詞・格助詞の少なさが6年生を分けている。4年生は名詞・接尾の少なさで5年生から分かれ、記号・読点の少なさで3年生から分かれる。

2) 考察

(1) 低学年教材に関する考察

1、2年生の説明文では、空白による分かち書きが多い。特に1年生の説明文では、文節単位で空白を入れる完全分かち書きであることを反映している。

(2) 高学年教材に関する考察

5年生の説明文に助動詞が少ないという結果に影響を及ぼしている要因のひとつは、敬体(です、ます)を使用しない説明文が最も多いことであると考

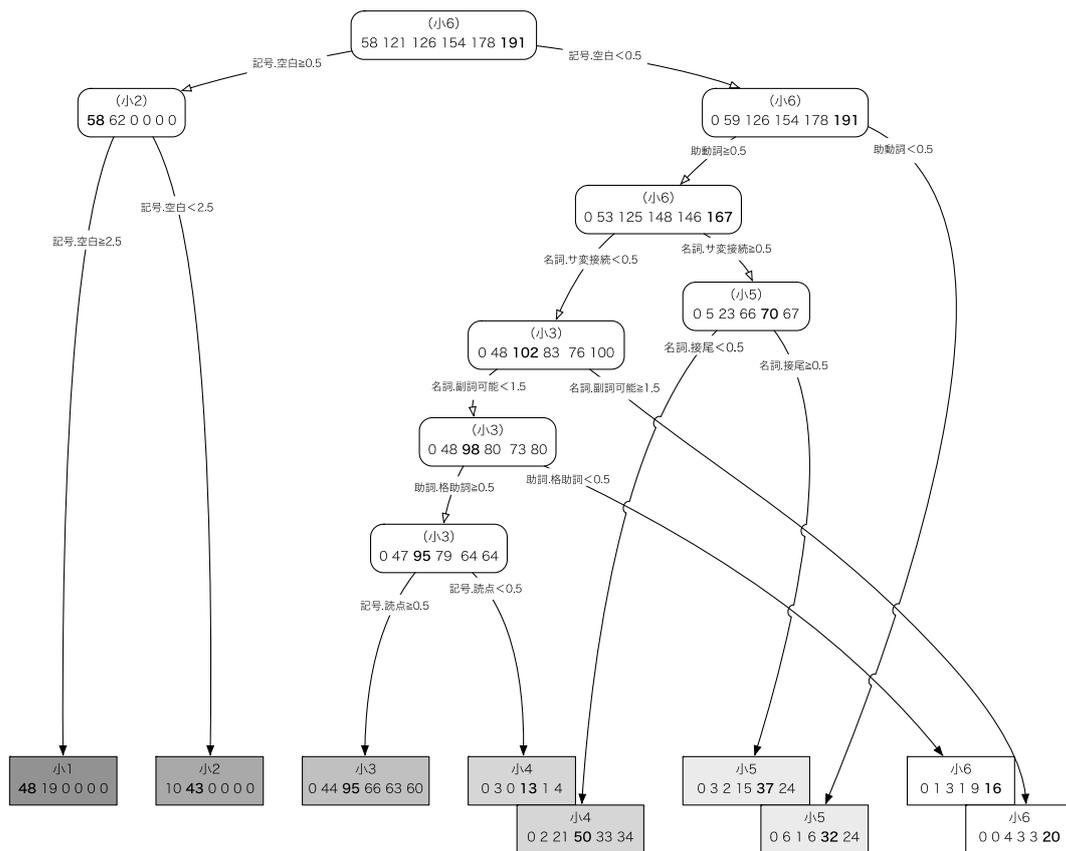


図5 学年を分類する樹形モデル

えられる。5年生におけるこうした文体の変化については、敬体の学習が中学年で取り扱われることとの関連性を指摘できる可能性もあるが、推測の域を出ない。また、漢語や外来語を使用する際に出現する品詞であるサ変接続名詞が5年生に多いことは、文章の難易度が高くなるということを示している。

さらに、5年生に多い名詞・接尾語は、単位等が使用されること、つまり説明内容に数値が使用されることの増加および派生語の増加が影響している。しかし、4年生の教材「ウナギのなぞを追って」でも同様に説明内容に数値が使用されていたが、数値を使用する文はすべて5年生と推定されていた。これは、5年生の教材「天気を予想する」においてグラフが使用され、グラフ内の数値を読み取りながら論が展開されることに起因している。そのため、5年生の教材として結果に影響を及ぼしたのは、数値の使用の増加よりも、派生語を構成する接尾辞の増加である可能性が高いといえる。

名詞・副詞可能は、曜日、月、量、割合、期間など、程度を表すような語を形成する品詞である。6年生にはある事象を説明する際に、何らかの指標やものさしを使用した論述が展開することに影響を受けていると考えられる。ただしこれには、時間をテーマにした教材「時計の時間と心の時間」のような個々

の教材が扱う題材の影響も否定はできない。格助詞の少なさが 6 年生を分けた要因についての考察は後述する。

(3) 中学年教材に関する考察

4 年生が記号・読点の少なさで 3 年生から分かれたことは、学年が上がることによって読点による文節の視覚的な明示が減少したことを表している。成人を対象とした漢字仮名混じり文の読みの眼球運動に関する一連の研究から、神部 (1998) は漢字や片仮名で示された主に内容を表す語で停留が起こることを示す。これはすなわち漢字仮名混じり表記が分節 (segmentation) 位置の指標となることを示している。それを踏まえると、漢字表記が増えることによって徐々に必要な読点が減るという影響関係を指摘できる。

3 年生は最後まで独立することがなく、突出した特徴が見いだせなかった。4 年生とあわせて考えると、上記のように特徴的な結果を見せた低学年と高学年に比べ、中学年の教材にはあまり特徴がない。それ自体が、低学年としての特徴を持った文章から徐々に高学年としての特徴を持つ文章へ移行する過渡期として、様々な特徴を少しずつ備えている学年としての特徴なのだと言えらることは可能である。

(4) 格助詞に関する考察と品詞の組み合わせの分析

格助詞の少なさが 6 年生を分けた要因を探るため、推定学年と実際の学年が一致した文を対象として形態素分解を行い、連続する品詞の組み合わせを抽出した。名詞・一般+助詞・格助詞は、3 年生 95 文中 129 個に対して、6 年生 36 文中 21 個と、36 文に期待される値 48.9 個より少なくなっている。抽出された格助詞には「という」「として」といった連語を形成するものと「が」「を」「に」など、主格・目的格を形成するものが含まれている。特に後者は、冒頭で述べた照応解決の中でも特にゼロ代名詞と関わりが深いと従前より指摘されてきた格助詞である (関ほか 2002; 桃内 2011)。

ゼロ代名詞の指示対象や生起条件は複雑で、その統計的解析・同定方法は確立されていないわけではない。しかし、名詞・一般+助詞・格助詞が 6 年生に少ないことと、結束性の高い文章の場合日本語では代名詞で明示されないことが多いことを考え合わせると、「が」「を」「に」の格助詞を伴って主格・目的格を形成する名詞が省略される、つまりゼロ代名詞を含む文が 6 年生に多くなっている可能性を指摘できる。

4. 総合考察

以上のように、本研究では、ある程度の学年間の違いと変容について判別することができた。学年差として抽出できた特徴には、係り受けの認識に必要な句や節の構造の把握を補助する空白や句読法 (読点)、照応解決に関連する代

名詞、句や節の構造の把握と関わるだけでなく、ゼロ代名詞の増加の間接的な証左とも考えられた格助詞が含まれていた。係り受けに関わる分節構造に関しては、まず、分かち書きによって視覚的・音声的に学習した後、読点として触れる。そして3年生から4年生への読点の減少から、漢字の増加による影響を指摘できる。読解において求められるものが、視覚的な情報に基づく句や節の境界の認識から、徐々に助詞などの文法的な標識に基づくものへと変化していくと考えられる。また、照応解決に関しては、学年が上がるほど代名詞が増加し、格助詞は6年生で減少することも示された。これらのことから、学年が上がるとともに読解時に照応解決を求められる機会が増え、また、ゼロ代名詞にも自然と触れる機会が増えるものと推測される。このように国語科の教科書の説明文は、徐々に多様な文型に触れる機会が増えるようになっており、これを通して照応解決等それぞれの構文の解析に熟達していくようになっている。

加えて、教材固有の特徴に起因して生じた可能性があるものを除いた特徴的な結果として、4、5、6年生でのサ変接続名詞の多さが挙げられる。これは代用動詞「する」を伴って動詞化するものであるが、和語系の動詞に加えてより多くの動詞のバリエーションを生む。このサ変接続名詞（動詞的名詞 Verbal Noun）を伴う動詞は、統語構造という点でも多様性をもっており（長谷川 1999, pp. 78-86）、例えば「麻子が戦後史（を/の）研究（φ/を）している」といった文の場合、丸括弧内のスラッシュの前後の要素のどちらとどちらを組み合わせた場合に自然な表現になるかを判断するには言語的な直観を要する。このようなことが文の構造の解析の容易さに関わるのかは不明だが、サ変接続名詞を用いた自然な表現を表出できるようになるためには、これを用いた多くの表現に触れることが必要だと考えられる。また、サ変接続名詞を伴う文の読解は、少なくとも、語彙的な多様さという点で高学年の学習における1つの課題と考えられる。

このように、小学校国語教科書における説明文は、読解力の基礎となる力の育ちに寄与するような特徴を有していると、部分的ではあるが指摘できる。冒頭で述べた「高次読解力」と比べて、こうした力は授業の中で明示的に取り立てて指導することは困難である。そうだとすれば、これらの力を身に付けるためには、何度も繰り返し多様な表現や文型に多く触れることが最も重要になる。そのことを踏まえると、小学校の国語教科書が徐々に多様な文型を扱うようになっていくという特性をもつことは一定の学習効果をもっていることを示している。しかし、これらは推測の域を出ておらず、さらにその品詞内の具体的な表現に踏み込んだ計量と分析が必要である。また、こうした教科書の特性が実際にどの程度子ども達の発達に即したものになっているのか、またこれによって国語科はどの程度構文解析の発達を支援しているのかについても、今後さ

らに検討する必要がある。

さらに、本研究では一社のみ国語科の教科書を分析対象としたため、教材に教科書会社特有の偏りがあった可能性もある。したがって、この樹形モデルが国語科の説明文教材すべてに適用できる普遍性をもつとはいきれない。また、分析対象とした教科書は平成20年告示版の学習指導要領に基づくものであり、平成29年の学習指導要領改訂に伴い教材にも削除・追加等の変更がなされている。それらを踏まえ、今後、複数の教科書会社による現行の国語教科書を対象としたさらなる検討が望まれる。

【注】

- 1) 投稿時の2022年における現行の小学校国語科教科書は(2020(令和2)年度版)である。現行の同社の教科書では、以下の教材が変更されている。2年「しかけカードのつくり方」4年「大きな力を出す」「動いて、考えて、また動く」5年「生き物は円柱形」「天気を予想する」6年「自然に学ぶ暮らし」が削除。2年「馬のおもちゃのつくりかた」4年「思いやりのデザイン」「世界にほこる和紙」5年「言葉の意味が分かること」「固有種が教えてくれること」6年「メディアと人間社会」が追加。

【参考・引用文献】

- 新井紀子(2018) AI vs. 教科書が読めない子どもたち 東洋経済新報 東京
- ARAI, N. H., TODO, N., ARAI, T., BUNJI, K., SUGAWARA, S., INUZUKA, M., MATSUZAKI, T., & OZAKI, K. (2017) Reading skill test to diagnose basic language skills in comparison to machines. Proceedings of the 39th Annual Cognitive Science Society Meeting 1556-1561.
- 長谷川信子(1999)生成日本語学入門 大修館書店 東京
- 船所武志(2014)国語科教育における論理的思考能力育成に関する一考察 -小学校の説明文・意見文教材を中心に-. 四天王寺大学紀要, 57 : 235-246
- 井上一郎(1986)説明文の語彙. 国語科教育, 33 : 139-147
- Ishida, M. (2017) RMeCab: interface to MeCab. R package version 0.99998.
- 石田栄美・安形 輝・野末道子・久野高志・池内 淳・上田修一(2004)文体からみた学術的文献の特徴分析. 2004年度三田図書館・情報学会研究大会発表論文集, 33-36
- 石井英真(2002)「改訂版タキソノミー」によるブルーム・タキソノミーの再構築-知識と認知過程の二次元構成の検討を中心に-. 教育法法学研究, 28 : 47-58
- 神部尚武(1998)日本語の読みと眼球運動. 荻阪直行(編)読み一脳と心の情

- 報処理. 朝倉書店 東京, 1-16
- 岸 学・綿井雅康・谷口淳一 (1989) 説明文の構造とその理解について—小学校国語教科書の分析に基づく検討—. 東京学芸大学紀要, 40 : 77-86
- 古賀洋一 (2020) 説明的文章の読解方略指導研究—条件的知識の育成に着目して. 溪水社 広島
- 国立情報学研究所 (2016) 文章を正確に読む力を科学的に測るテストを開発／産学連携で「読解力」向上を目指す研究を加速.
<http://www.nii.ac.jp/news/release/2016/0726.html> (Accessed 2022.3.11)
- KUDO, T., YAMAMOTO, K., & MATSUMOTO, Y. (2004). Applying Conditional Random Fields to Japanese Morphological Analysis. Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing, 230-237
- 間瀬茂夫 (2015) 高等学校における高次読解力の評価のあり方：読解力評価問題の活用. 国語教育研究, 56 : 219-230
- 間瀬茂夫 (2017) 説明的文章の読みの学力形成論. 溪水社 広島
- 光野公司郎 (2005) 論理的な文章における効果的な構成指導の方向性：論証の構造を基本とした新しい文章構成の在り方. 国語科教育, 57 : 60-67
- 桃内佳雄 (2011) 日本語文章におけるゼロ代名詞解析に関する基礎的考察, 北海学園大学工学部研究報告, 38 : 100-119
- 村上征勝 (2000) 文章の計量分析—その歴史と現状—. 計測と制御, 39:216-222
- 難波博孝・東広島市立木原小学校 (2009) 楽しく論理力が育つ国語科授業づくり. 明治図書 東京
- 大内善一 (2011) 9. 説明文. 日本国語教育学会 (編) 国語教育総合事典, 朝倉書店 東京, 292
- R CORE TEAM (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
<https://www.R-project.org/> (Accessed 2022.3.11)
- 佐藤敏紀・橋本泰一・奥村学 (2017) 単語分かち書き辞書 mecab-ipadic-NEologd の実装と情報検索における効果的な使用方法の検討. 言語処理学会第 23 回年次大会, NLP2017-B6-1
- 関和広・藤井敦・石川徹也 (2002) 確率モデルを用いた日本語ゼロ代名詞の照応解析. 自然言語処理, 9(3) : 63-85
- THERNEAU, T., ATKINSON, B., & RIPLEY, B. (2017) Rpart: Recursive partitioning and regression trees
<https://CRAN.R-project.org/package=rpart> (Accessed 2022.3.11)