

説明的文章の要点把握のための読解方法の有効性評価に用いる 正解文データセット (CAKeS) の作成

渡邊 幸佑 (兵庫教育大学) [†]

Correct Answer Key Sentences Dataset for Evaluating Reading Methods for Extracting Key Sentences of Explanatory Texts

Kosuke Watanabe (Hyogo University of Teacher Education)

要旨

国語科の「読むこと」の指導において、説明的文章の要点を文単位で把握する読解方法(キーセンテンス把握方法)が教えられている。これまで様々な方法が提案されてきた。しかし、そのような方法によりどの程度正確に文章の要点を把握できるか定量評価されていない。他方、情報科学の自動要約の研究では、コンピュータによる要約を定量評価する方法がある。その一つに、要点として本来抽出すべき重要な文(正解文)と、コンピュータが抽出した文(抽出文)の一致度を測るものがある。この評価を実施するためには、人手で正解文を定めておく必要がある。しかし、国語教科書の説明的文章を対象に正解文を定めたデータセットはない。そこで、説明的文章を対象に研究力者3名と正解文を設定した。どの文を正解文とするかについて研究協力者間での判断は概ね一致したため、今回定めた正解文は要点として本来抽出すべき重要な文とみなしてよいと考えられる。

1. はじめに

『中学校学習指導要領(平成29年告示)』の国語科第一学年の「読むこと」に、説明的文章を用いた指導事項として、「文章の中心的な部分と付加的な部分、事実と意見との関係などについて叙述を基に捉え、要旨を把握すること。」(p.31)とある。この事項に関して、『【国語編】中学校学習指導要領(平成29年告示)解説』では、「キーワードやキーセンテンスなどに留意して情報を整理し、正確に要旨を捉えられるようにすることが重要である。」(p.68)とある。

これまで要点を把握するための読解方法が様々提案されてきた(大西 1981, 樺島 1983, 佐久間 1989, 岡本 1990, 石黒 2009, 川西 2014, 井上 2014, 白石 2014, 渡部 2015, 岡田 2016, 渡邊 2023a など)。しかし、従来提案されてきた読解方法は互いに異なっている。従来提案されてきた読解方法によって文章の要点をどの程度正確に把握できるか複数の文章サンプルを用いた定量評価はされていない(渡邊 2023b)。そのため、どの方法が優れており、どの方法を生徒に教えるべきか明らかでないのである。真に有効な読解方法を生徒に教えるためには、読解方法の定量的な評価が必要である。

他方、コンピュータを用いて要約を自動的に作成する自動要約の研究では、コンピュータの作成した要約を定量的に評価する手法が蓄積されてきた。たとえば、文章中から要点として重要な文を抽出させること(重要文抽出)の評価方法として、要点として本来抽出すべき重要な文(以下、正解文)と、コンピュータが抽出した文(以下、抽出文)の一致度を測るものがある(奥村・難波 2005)。正解文は人手で設定しなければならない。そこでテキスト自

[†] watten6210@gmail.com

動要約の評価型プロジェクト TSC(Text Summarization Challenge)が実施され、正解文の設定がされてきた (Fukushima and Okumura 2001, Okumura et. al. 2003, Hirao et. al. 2004)。

しかし、従来の正解文の設定は、新聞記事を対象としたものである。つまり、国語の教科書の説明的文章を対象としたものではない。国語教科書の説明的文章の要点を把握するための読解方法の有効性を評価するためには、国語教科書の説明的文章の正解文を定めたデータセットが必要である。

そこで、本稿筆者は、中学校国語科の教科書の説明的文章を対象に、研究協力者と正解文を定める作業をし、説明的文章の正解文データセットを作成した。

なお、このデータセットを「CAKeS (ケイクス)」と命名する。Correct Answer Key Sentences の下線部分をとりだしたものである。

2. 正解文データセット(CAKeS)の作成方法

2.1 対象文章

令和3年度版の中学校国語教科書における説明的文章 48 編とする。光村図書、東京書籍、三省堂、教育出版の中学校第 1 学年から第 3 学年の教科書に収められている説明的文章である。

2.2 研究協力者

研究協力者 3 名で作業した。国語科教育学の研究者 1 名、日本語学の研究者 2 名 (うち 1 名は小学校教員経験者) である。

2.3 正解文の数

1 文、並びに、文章全体の文の数の 10%、30%、及び 50%とする (以下、1 文抽出、抽出率 10%などと言及する。)

テキスト自動要約の評価型プロジェクト TSC では、抽出率 10%、30%及び 50%の文抽出を行っており、1 文抽出は行っていない。一方、国語教科書の説明的文章において重要な 1 文を抽出できることが経験的に認められるため、本稿では 1 文抽出も含め検討した。

なお、この正解文の数は目安であり、実際には研究協力者との合議により若干の増減が生じている。

設定する正解文の数に幅を持たせている。これにより、要約の制限字数の多寡に応じた最適な文抽出方法を検討できる。たとえば、最も重要な 1 文を抽出するときと、文章全体を 50%程度に縮小するときとは、異なる読み方を採用すべきと予想されるためである。

2.4 作業の流れ

以下のように、まず個人作業を行い、その後全体作業を行った。

- 1) 個人作業：個人で正解文を設定する。このとき他の者の作業内容は見ない。
- 2) 全体作業：その後全体で個人の結果を突合し合議により全体の正解文を定める。

2.5 作業の期間・場所

作業完了まで 2022 年度の 1 年間を要した。教科書会社 1 社の 1 学年分の個人作業 (文章 3~5 編) を 30 日程度の期間で行い、月末に合議を行うことを目安とした。

研究協力者は互いに遠隔地に居住していたため、合議はオンライン会議システムを利用し行った。研究協力者のうち 1 人を司会とし合議を進行した。合議の際は、個人作業の時点で 3 者とも判断が一致しているところは確認程度で済ませ、3 者間で判断が分かれた文について集中的に議論した。

3. 正解文データセット(CAKeS)の諸情報

以上の作業を通じ得られた正解文データセット(CAKeS)は次のものである。なお、どの文を正解文としたかについては将来的に公表したい。

3.1 対象文章

対象文章のタイトル、教科書、発行者、文章全体の文の数、正解文の数は、表1のとおりである。

表1 対象文章

文章タイトル	教科書	発行者	文章全体の文の数	正解文の数(1文)	正解文の数(抽出率10%)	正解文の数(抽出率30%)	正解文の数(抽出率50%)
ダイコンは大きな根?	国語1	光村図書	34	1	3	11	17
ちょっと立ち止まって	国語1	光村図書	36	1	4	11	18
比喩で広がる言葉の世界	国語1	光村図書	29	1	4	9	14
「不便」の価値を見つめ直す	国語1	光村図書	79	1	8	24	40
「言葉」をもつ鳥、シジュウカラ	国語1	光村図書	63	1	6	19	32
モアイは語る	国語2	光村図書	79	1	8	24	40
クマゼミ増加の原因を探る	国語2	光村図書	91	1	9	27	47
君は「最後の晩餐」を知っているか	国語2	光村図書	99	1	10	31	52
人工知能との未来	国語3	光村図書	34	1	3	10	17
作られた『物語』を超えて	国語3	光村図書	71	1	7	21	36
人間と人工知能と創造性	国語3	光村図書	46	1	5	14	23
誰かの代わりに	国語3	光村図書	61	1	6	18	31
オオカミを見る目	新しい国語 1	東京書籍	55	1	6	17	28
ニュースの見方を考えよう	新しい国語 1	東京書籍	79	2	9	24	40
私のタンポポ研究	新しい国語 1	東京書籍	82	1	8	25	41
正しい言葉は信じられるか	新しい国語 2	東京書籍	72	1	7	22	36
サハラ砂漠の茶会	新しい国語 2	東京書籍	51	1	5	16	26
黄金の扇風機	新しい国語 2	東京書籍	50	1	5	15	25
ハトはなぜ首を振って歩くのか	新しい国語 2	東京書籍	99	1	10	30	50
絶滅の意味	新しい国語 3	東京書籍	85	1	9	26	43
幸福について	新しい国語 3	東京書籍	135	1	14	41	68
何のために「働く」のか	新しい国語 3	東京書籍	82	1	8	25	41
ペンギンの防寒着	現代の国語 1	三省堂	22	1	2	7	11
意味と意図	現代の国語 1	三省堂	44	1	4	13	22
クジラの飲み水	現代の国語 1	三省堂	45	1	5	14	23
玄関扉	現代の国語 1	三省堂	40	1	4	12	20
壁に残された伝言	現代の国語 2	三省堂	99	1	11	32	53
動物園でできること	現代の国語 2	三省堂	96	1	10	29	48
ジャンケン、なぜグー・チョキ・パーの三種類なのか	現代の国語 2	三省堂	35	1	4	11	18
一〇〇年後の水を守る	現代の国語 2	三省堂	86	1	9	26	43
人間は他の星に住むことができるのか	現代の国語 2	三省堂	56	1	6	17	26
俳句の世界	現代の国語 3	三省堂	46	1	5	14	23
フロン規制の物語	現代の国語 3	三省堂	88	1	9	26	44
間の文化	現代の国語 3	三省堂	42	1	4	13	21
「文殊の知恵」の時代	現代の国語 3	三省堂	41	1	4	12	21
言葉がつながる世界遺産	中学国語 1	教育出版	65	1	7	20	33
子どもの権利	中学国語 1	教育出版	41	1	4	12	20
自分の脳を知っていますか	中学国語 1	教育出版	61	1	6	18	30
森には魔法使いがいる	中学国語 1	教育出版	89	1	9	27	45
水の山 富士山	中学国語 2	教育出版	53	1	5	16	27
紙の建築	中学国語 2	教育出版	65	1	7	20	33
日本の花火の楽しみ	中学国語 2	教育出版	49	1	5	15	25
学ぶ力	中学国語 2	教育出版	90	1	9	27	45
ガイアの知性	中学国語 2	教育出版	65	1	7	20	33
問いかける言葉	中学国語 3	教育出版	52	1	5	16	26
async——同期しないこと	中学国語 3	教育出版	62	1	6	19	31
AIは哲学ができるか	中学国語 3	教育出版	31	1	3	9	16
青春の歌——無名性の光	中学国語 3	教育出版	73	1	7	22	37

3.2 文の数え方

対象文章に含まれる文の数え方は次のとおりである。

【原則】

句点、感嘆符、疑問符（以下、句点等）の数を数える。その数を文の数とみなす。ただし、タイトル中の句点等は数えない。

【例外 1】

鍵括弧の後に「と」などの引用の格助詞がある場合、鍵括弧中の句点等はないものとみなす。

（例 1）1 文と数えるもの

「おお、賛成と反対の人が同じくらいいるんだなあ。」と思いませんか。

（池上彰「ニュースの見方を考えよう」）

（例 2）1 文と数えるもの

例えば、インタビューを受けたご本人から、「私の話したことごとく一部しか伝えられなかった。これでは視聴者に誤解されかねない。」というような抗議が出ることも、ときにはあるのです。

（池上彰「ニュースの見方を考えよう」）

（例 3）2 文と数えるもの

「子どもは黙って大人の言うことを聞きなさい。」こんなことを言われて理不尽な思いをしたことはありませんか。

（大谷美紀子「子どもの権利」）

【例外 2】

- 「比喻で広がる言葉の世界」の中の詩は文として数えない。
- 「青春の歌—無名性の光」の中の短歌とその詠み手はまとめて 1 文と数える。

3.3 一致率

複数人の読み手が 1 つのテキストに対して何を重要視するかについて必ずしも一致するとは限らない(浅原ら 2015)。研究協力者の個人作業の結果が互いにどの程度一致するかをまとめたものが表 2 である。

表 2 一致率

	[3,0]	[2,1]	[1,2]	[0,3]	一致率
1文抽出	48	1	2	3008	0.941
抽出率10%	282	22	35	2720	0.832
抽出率30%	831	88	96	2044	0.819
抽出率50%	1375	156	178	1350	0.805

[3,0] : 3 名とも重要であると判断した文の頻度

[2,1] : 2 名が重要であると、1 名が重要でないと判断した文の頻度

[1,2] : 1 名が重要であると、2 名が重要でないと判断した文の頻度

[0,3] : 3 名とも重要でないと判断した文の頻度

一致率 = $[3,0] \div ([3,0] + [2,1] + [1,2])$

なお、上記の式で定義する一致率の分母及び分子に[0,3]は含めていない。一致率の分母及び

分子に[0,3]を含めると正解文判断一致の実状を表せないからである。たとえば、100文からなる文章から1文を正解文として抽出する場合であって、3名とも異なる文を重要だと判断したとする。この場合、一致率の分母及び分子に[0,3]を含めて、一致率※ = $([3,0]+[0,3]) \div ([3,0]+[2,1]+[1,2]+[0,3])$ と考えると、

$$\text{一致率※} = (0+97) \div (0+0+3+97) = 0.97$$

というように、3者の正解文判断が不一致にもかかわらず、一致率※は極めて高く、実状に合わないためである。

表2の一致率から、1文抽出の場合、研究協力者の判断は94%ほど一致している。抽出率10%の場合、研究協力者の判断はおおむね83%ほど一致している。抽出率30%の場合、研究協力者の判断は82%ほど一致している。抽出率50%の場合、研究協力者の判断は81%ほど一致している。このように、個人作業での研究協力者の判断は概ね一致している。さらにその後、合議により十分な議論を重ね、各人が納得した上で、全体での正解文を定めている。そのため、合議により定めた正解文は、要点として本来抽出すべき重要な文とみなしてよいと考えられる。

また、表2から、一致率は正解文の数が増加するにつれ減少する。ここから、制限字数の少ない要約であるほど、読み手間の判断が一致しやすいことが示唆される。

4. 読解方法の有効性評価の方法

今回作成した正解文データセット (CAKeS) を用いて、説明的文章のキーセンテンスを抽出する読解方法を対象に、その有効性を評価できる。

評価尺度は、以下の再現率、精度、F値を用いる。F値が高いほど要点を捉える読解方法として有効性が高いと判断する。

$$\text{再現率} = a \div b$$

$$\text{精度} = a \div c$$

$$\text{F値} = 2 \times \text{再現率} \times \text{精度} \div (\text{再現率} + \text{精度})$$

a = ある読解方法で抽出した文に含まれる正解文の数

b = 正解文の数

c = ある読解方法で抽出した文の数

再現率は、正解文をどれだけ網羅的に拾っているか測る尺度である。精度は、正解文のみを拾っているか（不要な文を拾っていないか）測る尺度である。F値は、再現率と精度の調和平均であり、再現率と精度を総合評価するために用いる。F値が高いほど要点把握のために有効性が高い、と判断する。

5. 今後の課題

今後の課題は、様々な読解方法の有効性評価、正解文データセット (CAKeS) のオープン化、正解文データセットの大規模化、正しい読みの存否の検討である。

これまで様々な読解方法が提案されてきた。「しかし」や「このように」などの特定の接

続表現を含む文を抽出する方法、内容のまとめの文や問題提起と答えを述べる文を抽出する方法、事実と意見に着目し意見を抽出する方法などである。こうした読解方法の有効性を評価し、真に有効な読解方法を見出せると期待される。なお、事実と意見に着目し意見を抽出する方法については既に有効性評価がされている(渡邊 2023c,2023d)。また、新しく提案した読解方法(渡邊 2023a)についても有効性評価を行いたい。

正解文データセット(CAKeS)を将来的にはオープン化し、あらゆる研究者の利用に供したい。その際、著作権、利用しやすさ等に留意する。なお、オープン化以前に正解文データセット(CAKeS)を利用したい方にとっては個別に相談されたい。

今回の正解文データセット(CAKeS)は、研究協力者3名で作業した。もし膨大な数の研究協力者とともに正解文データセットを作成できるならば、読みの実態がより鮮明に浮かび上がる。人々はどの文をキーセンテンスだと思うか、どのような要約を理想とするか。こうしたデータを全国・世界規模で集約することで、人間の文章理解の普遍性と独自性を明らかにしたい。

このような正解文データセットの大規模化が実現すれば、正しい読みの存否を検討できる。多くの人の間で一致する読み方を正解とみなすことができる。一方、多くの人の間で一致する読み方などないかもしれない。読みに正解がないとすれば、国語科教育はどうあるべきか。データ駆動型の国語科教育学は、従来の国語科教育の在り方を抜本的に変革する可能性を秘めている。

謝 辞

本稿は公益財団法人日本漢字能力検定協会、漢字・日本語教育研究助成「説明的文章の要点を把握するための読解方法の開発・評価」により実施した研究成果の一部である。

文 献

- 浅原正幸・杉真緒・柳野祥子.(2015)「BCCWJ-SUMM:『現代日本語書き言葉均衡コーパス』を元文書とした要約文書コーパス」『第7回コーパス日本語学ワークショップ予稿集』,pp.285-292.
- 石黒圭(2009).『よくわかる文章表現の技術Ⅱ 文章構成編 [新版]』明治書院.
- 井上善弘(2014).「接続語や文末からキーセンテンスを探すヒント指導」『教育科学国語教育』,No.778 pp.70-71.
- 大西忠治(1981).『説明的文章の読み方指導』明治書院.
- 岡田圭亮(2016).「中学校の説明的文章教材における要約指導の研究:初読後の文章把握を目的とした要約の方法」『上越教育大学国語研究』,30, pp42-55.
- 岡本明人(1990).「要約の力」がつく指導のヒント「要約」とは何か」『教育科学国語教育』, 32:5, pp.13-16.
- 奥村学・難波英嗣(2005).『テキスト自動要約』オーム社.
- 樺島忠夫(1983).「文章構造」『朝倉日本語新講座5 運用Ⅰ』朝倉書店,pp.118-57.
- 川西良治(2014).「段落ごとにキーセンテンスを探すヒント指導」『教育科学国語教育』, No.778, pp.68-69.
- 佐久間まゆみ(1989).「文章の統括と要約文の構造特性」『文章構造と要約文の諸相』, ころしお出版, pp.184-228.
- 白石範孝(2014).「要約指導の方法 白石流 『問い』と『答え』と『要旨』の三点から

の要約指導」『教育科学国語教育』,No.778, pp.24-27.

Takahiro Fukusima and Manabu Okumura(2001) “Text Summarization Challenge: Text Summarization Evaluation at NTCIR Workshop2”, *Proceedings of the Second NTCIR Workshop on Research in Chinese & Japanese Text Retrieval and Text Summarization*,pp.45-50.

Tsutomu Hirao, Manabu Okumura, Takahiro Fukusima and Hidetsugu Nanba(2004) “Text Summarization Challenge 3 - Text Summarization Evaluation at NTCIR Workshop 4”, *Proceedings of the Fourth NTCIR Workshop*,pp.407-411.

Manabu Okumura, Takahiro Fukusima and Hidetsugu Nanba(2003) “Text summarization challenge 2: text summarization evaluation at NTCIR workshop 3”, *Proceedings of the HLT-NAACL 03 on Text summarization workshop* ,Vol.5, pp.49-56.

渡邊幸佑(2023a). 「説明的文章の要点把握のための読解方法—文脈効果に基づく文の重要度判定—」『読書科学』,64:2,pp.84-94.

渡邊幸佑(2023b). 「説明的文章の要点を把握するための読解方法についての批判的検討」『上越教育大学国語研究』,37,pp.27-41.

渡邊幸佑(2023c). 「「事実」と「意見」の再定義」『読書科学』,64:1,pp.1-14.

渡邊幸佑(2023d). 「事実と意見に着目した読み方により説明的文章の要点をどの程度正確に捉えられるか」『解釈』,69:5・6,pp.2-11.

渡部洋一郎(2015). 「授業の導入部における要約指導の効果」『教育科学国語教育』,No.792, pp. 9-13.

関連 URL

NTCIR TSC <http://research.nii.ac.jp/ntcir/index-ja.html>