

# 小学校での NIE 支援に向けた新聞記事からの テンプレートによる問題文生成の検討

國方海秀† 安藤一秋‡

香川大学大学院創発科学研究科† 香川大学創造工学部‡

## 1. はじめに

近年、小学校や中学校などの義務教育機関において、Newspaper In Education, 通称 NIE が実践されている。日本新聞協会や経済協力開発機構の調査[1]によると、NIE を実践することで児童の読解力や学力の向上に効果があることが示されている。しかし、NIE を実践するにあたり、「カリキュラムとの調整が難しい」や「教材準備の時間が足りない」など、NIE を実践する教師の負担が大きくなるのが課題になっている。

いくつかの新聞社は、NIE を実践する教師を支援するために NIE ワークシートを配信している。NIE ワークシートは、新聞記事とその記事に関連する数個の問題で構成される。しかし、ワークシートを配信している新聞社は 2023 年 1 月時点で 17 社しかない。また、小学校での NIE 実践では、地域に関連する記事が採用される傾向があるが、該当地域の新聞社がワークシートを配信していない場合や、配信していても教師が児童に触れてほしいテーマが存在していない場合、教師が独自にワークシートを作成する必要がある。

そこで、本研究では、NIE を実践する小学校教師を支援するため、NIE ワークシートの自動生成を目的とする。本稿では、事前検討として、テンプレート形式で問題文を生成する手法の実現可能性について検討する。

## 2. 我々の先行研究の概要

我々は、NIE ワークシートの自動生成に向け、問題文分析を進めてきた。以下、我々の先行研究[2]で実施した分析結果の概要を示す。

### 2.1 出題方式の分析

まず、NIE ワークシートに採用される問題文の出題方式に着目し、傾向の有無を確認した[2]。地方新聞社 9 社から 210 件のワークシートを任意選択し、各ワークシートに採用された問題文 708 件を分析対象とした。

NIE ワークシートに採用される問題文がどのような形式で出題されるかを人手で確認し、一般的な出題方式に分類した結果、全 17 ジャンルに分類できた。その中でも文章で答える問題が 369 件、単語で答える問題が 265 件と多く、全体の約 74% を占めていることがわかった。他の出題方式では、文章の空所を埋める穴埋め問題が 9% (79 件) や回答者に調査を促す調査問題 8% (72 件) が見られたが、すべて 80 件未満となった。

### 2.2 問題文の分析

問題文の自動生成手法としては、問題文の構造を利用してテンプレートで生成する手法と汎用言語モデルを利用して文生成する手法が考えられる。我々の事前調査において、小学生向けの問題文はシンプルな構造の問題も多いことを確認したことから、まずはテンプレートによる問題文の自動生成について検討する。

テンプレートによる問題文の自動生成を実現するためには、問題文に構造的な特徴の存在が必要となる。我々の先行研究[2]では、疑問詞に着目した問題文の分析と文末の表現に着目した問題文の分析を実施した。以下、その概要について述べる。

#### 2.2.1 5W1H の分析

問題文を構成する要素の 1 つとして疑問詞に着目し、ワークシートに採用される問題文 708 件を人手で確認した。文中の疑問詞を一般的な疑問詞とされる 5W1H に分類し、その件数を調査した結果、「何が」や「何を」といった what の要素を含む問題文が 34.6% (245 件)、「どんな」や「どのように」といった how の要素を問う問題が 28.2% (200 件) 存在し、その他の 5W1H と比べて明確に差があることを確認した。

また、how 要素に関しては、「どのような」や「どんな」といった多様性があるため、その内訳を調査した。その結果、「どのような」が 71 件、「どんな」が 80 件と、連体修飾するタイプの表現が全体の約 76% を占めることを確認した。

#### 2.2.2 文末の分析

問題文の構造的特徴は、文末にも表れる可能性があると考え、問題文を文節に分けて文末部分を人手で確認した。その際、名詞を N、数字を 0 に一般化し、「書きましょう」や「書いてください」など意味が等しいものは一意に集約した。

その結果を表 1 に示す。表中の割合は、分析した問題文 708 文に対する割合である。表 1 より、解答を促す「書きましょう」、名詞を問う「N ですか」、調査問題に多く見られた「調べてみよう」、what 要素を問う文末「何ですか」が上位になり、その他の 5W1H は 2% 以下と低くなった。また、その他のうち、1 件のみ見られたものが 121 件存在し、問題文の文末における多様性を確認した。

#### 2.2.3 5W1H と文末表現の共起性の分析

次に 5W1H と文末表現の共起性に着目し、構造的な傾向の有無を確認した。表 1 の上位 4 件の文末表現に対し、それぞれの 5W1H の件数を人手で確認した結果を表 2 に示す。

Consideration of Question Sentence Generation from Newspaper Articles using Templates for NIE Support in Elementary Schools

†Kaishu Kunikata, Graduate School of Science for Creative Emergence, Kagawa University

‡Kazuaki Ando, Faculty of Engineering and Design, Kagawa University

表1 問題文における文末表現の出現割合

問題 text	件数	割合
書きましょう	151	21%
Nですか	71	10%
調べてみよう	67	9%
何ですか	57	8%
挙げましょう	16	2%
ありますか	14	2%
考えてみましょう	12	2%
その他	320	45%

表2 文末表現と5W1Hの共起割合

	what	who	when	where	why	how	合計	割合
書きましょう	11	3	0	1	0	32	47	31%
Nですか	8	3	2	6	0	42	61	86%
調べてみよう	1	0	1	2	0	11	15	22%
何ですか	2	0	1	1	0	3	7	12%
合計	22	6	4	10	0	88	130	38%

表3 テンプレートで生成可能な短文問題数

テンプレート	件数	割合
～は何ですか	24	4.5%
～は誰ですか	2	0.4%
～はどこですか	13	2.4%
～はいつですか	14	2.6%
～はなぜですか	7	1.3%
～はどのくらいですか	4	0.7%
～はどんなNですか	15	2.8%
～について調べてみよう	30	5.6%
合計	109	20.2%

表2より、最も割合が高くなった文末は「N ですか」で、その86%に5W1H表現が含まれることを確認した。その他は「書きましょう」、「調べてみよう」、「何ですか」間で10%程度の差が見られた。しかし、各々の割合は高くないため、他の視点で特徴を探る必要がある。

### 3. テンプレートによる問題文生成の検討

我わらの先行研究[2]の分析結果を踏まえて、テンプレートで問題文が生成できる可能性について検討する。表2において割合が最も高いhow要素と文末表現「N ですか」を組み合わせた問題文、表1において文末表現の出現割合が3位の「調べてみよう」問題文、そして、「いつですか」や「どこですか」などの5W1Hを文末として問う問題文のうち、短文で構成されている問題文、計539件を対象に、テンプレートで問題文が生成できる問題数を人手で確認する。

テンプレートで生成できると判断した件数を表3に示す。表3より、テンプレートで生成できる件数が最も多いものは「～について調べてみよう」であった。出題方式の分析で対象とした調査問題(72件)のうち約42%が生成可能であるといえる。次に、5W1H要素を文末に持つテンプレートにおいては、表1の文末表現にも多く見

られた「～は何ですか」が全単文に対して4.5%とあまり多く出現しなかった。5W1Hと文末表現の共起性の分析において、最多件数であったhow要素と文末「N ですか」の組み合わせ「～はどんなNですか」は15件であり、how要素を問う問題全体の12.5%を占めていた。

「～は誰ですか」、「～はなぜですか」、「～はどれくらいですか」はそれぞれ10件以下となったが、テンプレートで生成された問題文はシンプルであり、児童にとっても出題側の意図が通じやすいといえるため、少ない件数でも汎用性があればテンプレートに採用できるといえる。最終的に、テンプレートで生成可能な問題文は、合計109件のNIEワークシートで採用されており、単文で構成される問題文の約20%が自動生成可能と考える。

テンプレートで問題文生成が難しい例として、複数の文で構成される問題(例: 鍍金は100年たっても鮮やかな色が残ります。その理由を記事から探してください。[3])、テンプレートの穴埋め部分が長い問題(例: プロ野球のドラフト会議で1位指名された風間投手が「最高のプレゼントに感謝したい」と語った理由は何ですか。[4])などが挙げられる。これらの問題文ではテンプレートではなく、汎用言語モデルを用いた文生成による手法を検討する。

### 4. おわりに

本稿では、NIEワークシートの自動生成に向け、NIEワークシートとそれに付随する問題文の分析結果の概要を述べた後、テンプレートを用いた問題文の自動生成の可能性について検討した。

how要素と文末表現「N ですか」を組み合わせた問題文、「調べてみよう」タイプの問題文、そして、文末で5W1Hを問う問題のうち、短文で構成されている問題文について、テンプレートを用いた問題文の自動生成について実現可能性を検討した結果、単文で構成される問題文の約20%が自動生成可能であることを確認した。

今後は、分析を継続し、新しいテンプレートについて検討する。そして、テンプレートに当てはまる固有表現などの選出方法について検討する。その後、汎用言語モデルを用いた問題文の自動生成手法について検討し、NIEワークシートの自動生成を目指す。

### 謝辞

本研究の一部は、JSPS 科研費 19K12271 と 22K12294 の助成を受けて実施した。

### 参考文献

- [1] NIE 教育に新聞を, <https://nie.jp/>
- [2] 國方海秀, 安藤一秋, “NIE ワークシートの自動生成に向けた問題文の分析”, 第21回情報科学技術フォーラム(FIT2022), 2022.
- [3] 大分合同新聞, [http://www2.oita-press.co.jp/nie/img\\_file/worksheet/163825655623834-1.pdf](http://www2.oita-press.co.jp/nie/img_file/worksheet/163825655623834-1.pdf)
- [4] 山梨日日新聞, [https://www2.sannichi.co.jp/NIE/wp\\_nie/wp-content/uploads/2021/10/20211021\\_10\\_【小学校：高学年】-風間球打選手（2021年10月21日週刊こびと10面）.pdf](https://www2.sannichi.co.jp/NIE/wp_nie/wp-content/uploads/2021/10/20211021_10_【小学校：高学年】-風間球打選手（2021年10月21日週刊こびと10面）.pdf)