ことばつなぎゲーム: ゲーミフィケーションによる小学生の作文教育

大村 和正　　久保 圭　　黒橋 禎夫
京都大学大学院情報学研究科
omura@nlp.ist.i.kyoto-u.ac.jp　kaykubo.ktu@gmail.com　kuro@i.kyoto-u.ac.jp

1 はじめに

多くの小学生は作文に苦手意識を持っている [1]. この一因として，日本の作文教育が，読書感想文や生活作文など，自身の情緒を自由に記述させる作文課題を主としていることが挙げられる．課題という形では課題をこなす以上の意欲を持つことが難しく，自由記述式であるために何を書こうかと頭を悩ませることも多い．加えて，作文に対するフィードバックが少ないために，改善点が分からないまま苦手意識だけが募るという悪循環に陥ってしまう．

このような現状を改善するためには，文を楽しく組み立てられ，その作文に対するフィードバックを得られるような枠組みが必要であると考えられる．しかし，作文は，算数の計算問題などとは異なり，自動処理による評価が難しいという問題があった．

近年，大規模な言語資源が構築されるようになり，それらは自然言語処理の発展を支えている．構築された言語資源の中には，作文問題の作問・評価に十分利用可能なものもあり，前述の枠組みが現実味を帯びてきた．

そこで，本研究では，大規模な言語資源を利用し，小学生向けの作文ゲームを考案した． 言葉を繋げて単文・複文を作り，その結果にフィードバックを与えることができるゲームを提案する．以下では，このゲームを**ことばつなぎゲーム**と呼ぶ．

ことばつなぎゲームは，図 1 のように，語と語を格助詞で繋いで文を，文と文を接続助詞で繋いで蓋然的関係[1)]のある複文（以降，**関係**と呼ぶ) を作るゲームである．プレイヤーは，与えられた言葉カードをマスの形に合わせて配置し，助詞マークで繋いで文および関係を作る．

作成した文および関係は自動評価が可能であり，フィードバックを得ることができる．例えば，作った文が既存の言語資源にデータとして存在しなければ，図 2 のように「より適切な格助詞がある」などのフィードバックを返す．プレイヤーは配置を変えて再評価することができるため，作文課題には無いインタラクティブ性を持つ．

まとめると，ことばつなぎゲームは，文を組み立てるという作文時の作業のゲーミフィケーションである．言葉を繋ぐという作業を通して，コロケーション・助詞の用法・基本的な文の間の蓋然的関係など，作文に必要な知識を養うことができ，これによって，語彙力・思考力の育成に資することを目的としている．

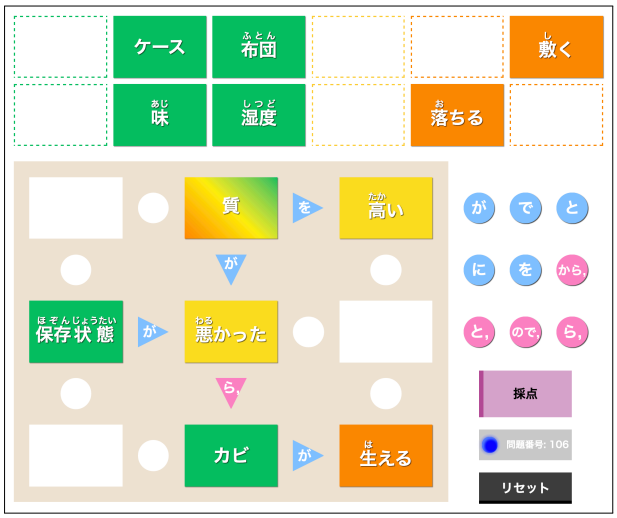

図 1　ことばつなぎゲームのデモ.

作ったもの		評価
作った関係	保存状態が悪かったら,カビが生える	★★★
	質が悪かったら,カビが生える	AIにとって新しい関係です
作った文	保存状態が悪かった	★★
	カビが生える	★★
	質が悪かった	★
	質を高い	より適切な格助詞があります(を→が)
		★8

図 2　図 1 の入力を自動評価した結果.

1)　ある事柄がある程度起こりうる/真であるという関係.

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

2　言葉カードセットの構築

2.1　問題の仕様

各問題では，12 枚の**言葉カード**と無制限に使用できる 9 個の**助詞マーク**が与えられる．これらをもとに文および関係を作る．

言葉カードの内訳は，名詞カードが 6 枚，動詞カード/形容詞カードが 5 枚，自由に単語を入力できるワイルドカードが 1 枚となっている．名詞・形容詞・動詞をそれぞれ緑色・黄色・橙色で表現するため，品詞という概念が難しくとも色を手掛かりに区別できる．単語の表記は単語難易度データベース [2] の頻出表記を利用し，読めない漢字がないようにルビが振られる．また，作文の自由度を上げるため，述語の活用は現在形・過去形・否定形・過去否定形のいずれかに変更できる．

助詞マークは，主要な格助詞 (が・を・に・で・と) が書かれたマーク 5 個と，蓋然的関係を表す接続助詞 (から・と・ので・ら) が書かれたマーク 4 個から成る．言葉カードと同様に，色によって助詞の役割を区別できる．マスに当てはめると形が三角形に変わり，三角の向きが係り受けの方向を表す．なお，格助詞「と」は並列助詞として使うこともできる．

2.2　蓋然的基本イベントペアの利用

文を組み立てる面白さを損なわないためにも，言葉カードセットから複数の関係文を作れることが望ましい．このために，本研究では，蓋然的基本イベントペア [3] に含まれる基本イベントの組 (以降，**基本イベントペア**と呼ぶ) を利用する．まず，これらの用語の説明を次に示す．

基本イベント　テキストから抽出した述語項構造をクラスタリングし，その中の高頻度なものを核とする表現．具体的には，格フレーム [4] から高頻度の述語項構造を基本イベントとして獲得する．

蓋然的基本イベントペア　蓋然的関係を持ち，前件と後件が共に基本イベントを含むイベントペア．例えば，「やっぱりお腹が空く → 駅でたぬきそばを食べる」が該当する．この例に含まれる基本イベントペアは，「お腹が空く → そばを食べる」である．

言葉カードセットは基本イベントペアから自動生成する．例えば，次のような基本イベントペアを考える．

	高依存	低依存
1 基本句	修飾語を付与して使う	修飾語を省略して使う
2 基本句以上	この項を含む基本イベントペアを使わない	
修飾なし	そのまま	

表 1　項の修飾語への依存度と修飾語の基本句数に応じた修飾語の扱いを整理した表．

(1)
a. 保存状態が悪い → カビが生える
b. 保存状態が悪い → 味が落ちる
c. 湿度が高い → カビが生える
d. 布団を敷く → カビが生える
e. カビが生える → ケースに入れる

これらの基本イベントペアを項と述語に分解すると，図 1 の言葉カードセットが得られる．このように，5 つの基本イベントペアをもとにするため，複数の関係文を作ることができる．

2.3　修飾語が重要な項への対処

蓋然的基本イベントペアは，2.2 節の例のように，基本イベントの他に修飾語や任意格を含みうる．そのままでは応用が難しいため，核にあたる基本イベントペアを利用している．しかし，修飾語が重要であるために基本イベントペアには蓋然的関係が認められない場合がある．例えば，次のようなものが挙げられる (下線は修飾語を表す)．

(2)
a. 線路状態が悪い → 左右に揺れる
b. 土曜日は満員の場合が多い
→ お早目のご予約が必要だ

このようなペアは修飾語を省略すると解釈が困難である．そのため，言葉カードセットの生成元に利用する場合は修飾語を補う必要がある．

この問題に対処するために，項の修飾語への依存度を定量化し，この依存度と修飾語の基本句[2)]数に応じて修飾語を自動処理する (表 1)．項の修飾語への依存度は，1) コーパス中で修飾語を伴う割合と，2) 抽象度の 2 つの軸で定量化する．

修飾語を伴う割合は，蓋然的基本イベントペア約 10 万組から算出する．抽象度については日本語抽象度辞書[3)]を利用する．日本語抽象度辞書は，Twitter での利用頻度上位 2 万単語 [5] の内，品詞が名詞・形容詞・動詞であるものを対象に，クラウド

2)　1 個の内容語および 0 個以上の付属語と定義される単位．
3)　http://sociocom.jp/~data/2019-AWD-J/

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

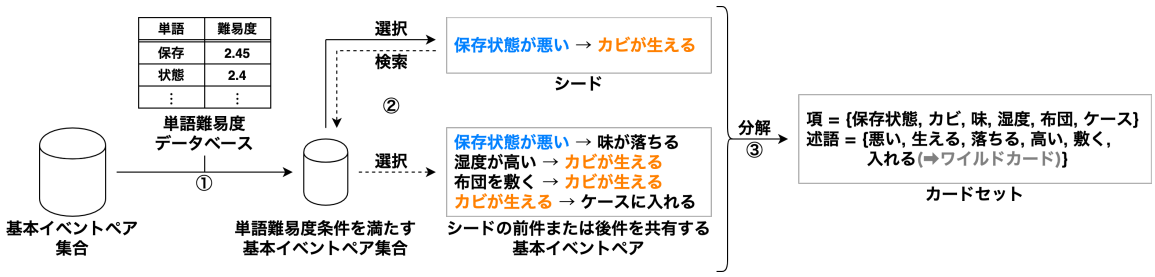

図 3　カードセットの生成手法の概要図.

ソーシングで 1 から 5 までの抽象度を付与したものである.

本研究では，次の条件を満たす時，その項は修飾語への依存度が高いとする.

[修飾語を伴う割合] × 5 + [抽象度] >= 5.75

これをもとに，修飾語の付与および基本イベントペアの除外を行なう.

2.4　言葉カードセットの生成手法

カードセットの生成手順は，次の 3 ステップから成る (図 3).

1. 単語難易度による閾値条件を設定し，条件を満たす基本イベントペアに対象を絞る.
2. ベースとなる基本イベントペアを 1 つ選び，これをもとにその他の基本イベントペアを 4 つ選択する.
3. 獲得された 5 つの基本イベントペアを項と述語に分解することで，カードセットを生成する.

以下では，各ステップについて説明する.

2.4.1　単語難易度による基本イベントペアの選択

対象が小学生であることを考慮し，単語難易度による語彙の調整を行う．このために，単語難易度データベース [2] を利用する.

単語難易度データベースは，形態素解析器 Juman++[6] の辞書に登録されている約 26,000 語を対象に，クラウドソーシングで習得時期を付与したものである．習得時期は 1 から 5 までの数値で表され，これを単語難易度として利用する．習得時期と数値の対応関係は，(1, 2, 3, 4, 5) = (小学生以前, 小学校低学年, 小学校高学年, 中学生以降, 単語を見聞きしたことがない) となっている.

本研究では，次の 3 つの閾値条件を設ける.

やさしい　基本イベントペアに含まれる各単語の単語難易度が全て 2.0 を超えない.

ふつう　基本イベントペアに含まれる単語の平均単語難易度が 1.5 以上かつ各単語の単語難易度が全て 2.5 を超えない.

むずかしい　基本イベントペアに含まれる単語の平均単語難易度が 2.0 以上である.

上限は厳密に超えないように定めるが，下限は平均単語難易度を用いることで多少のぶれを許容する.

やさしい	ふつう	むずかしい
152	1,363	797

表 2　カードセットの生成結果.

2.4.2　基本イベントペアの選択

単語難易度による閾値条件を満たす基本イベントペア集合から，カードセットの生成元となる 5 つの基本イベントペアを選択する．具体的には，ベースとなる基本イベントペア (以降，**シード**と呼ぶ) を 1 つ決め，シードの前件または後件を共有する基本イベントペアを無作為に 4 つ選択する．シードを含め計 5 つの基本イベントペアを獲得できなかった場合，カードセットを生成しない.

2.4.3　カードセットの生成

前段階で選択した 5 つの基本イベントペアを，項と述語に分解する．重複を除いた結果，項が 6 個かつ述語が 5 個または 6 個得られた場合，それをカードセットとする．述語が 6 個得られた場合は，シードに含まれない述語を無作為に 1 つワイルドカードに置き換える.

2.5　生成結果

2.4 節の手法をもとに，難易度ごとに言葉カードセットを生成した結果を表 2 に示す．個人が遊ぶ分には十分な規模のカードセットを生成できている.

また，修飾語を付加することで改善された例を図 4 に示す．修飾語によって他の単語と結びつきやすくなると同時に，解釈が難しい基本イベントペアを除いているため，文が作りやすくなっている.

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

図 4　修飾語を考慮することで改善される例．赤枠の部分が，付加された修飾語である．

3　入力の自動評価

プレイヤーが盤面のデータを送信すると，その入力が自動評価される．この自動評価は，1) 文および関係の認識，2) 文および関係の自動評価，3) 評価結果のフィードバックという手順で行われる．

3.1　文および関係の認識

格助詞で繋がれる経路の内，項のカードから始まり述語のカードで終わるものを文とみなす．また，認識された文に対し，接続助詞で繋がれる 2 文を関係とみなす．接続助詞について，接続元は述語のカードでなければならないが，接続先は項と述語のどちらでも良いとする．図 1 のように，接続元/接続先の文が複数存在する場合は，全ての文の組み合わせを考慮する．

3.2　文および関係の自動評価

文の評価方針として，項と述語の共起頻度が高く (=慣用的であり)，より系列長が長い文を高く評価する．文の評価は，格フレーム[4]の用例数をもとに決定する．具体的には，用例がある格フレーム cf に対し，項と格の組 (a, c) ごとに次のようなスコア関数 S でスコアを計算する．

$$S(a,c,cf) = 0.5 + 0.5 \times min(1, \frac{f_{a,c,cf}}{f_{cf}} \times 5)$$

$f_{a,c,cf}$: 格 c における項 a の頻度

f_{cf} : 格フレーム cf の頻度

例えば，「筍が顔を出す」という文のスコアは次のように計算される．

$S($ 筍, が, 出す$_6) + S($ 顔, を, 出す$_6) = 0.5 + 1 = 1.5$ [5]

なお，複数の格フレームに用例がある場合，スコアが最大となるものを選ぶ．

4)　京都大学格フレーム (https://www.gsk.or.jp/catalog/gsk2018-b) を利用する．

5)　「出す$_i$」は，述語「出す」の i 番目の格フレームを表す．

スコアをそのまま文の評価とすると解釈が難しいため，文の評価 e はスコア s をもとに次のように決定する．

$$e = \begin{cases} ☆☆ & (s >= 1) \\ ☆ & (0 < s < 1) \\ ? & (s = 0) \end{cases}$$

関係の評価は，基本イベントペアデータのいずれかを包含すれば “☆☆☆”，そうでなければ “？” とする．包含関係を調べる際は，述語の否定の極性まで一致することを確認する．

3.3　評価結果のフィードバック

評価が “？” であるものは，その理由に応じて次のような説明文を返す．

- 文が不完全である (項または述語が欠けている，述語から項に係っているなど)
- より適切な格助詞がある
- 未知の文/関係である

3 点目については，誤り報告機能を付けることで未知の関係文の収集が可能になると考えられる．

4　関連研究

日本語の作文に関する実応用に向けた取り組みは，文法誤り訂正 [7] や記述式答案の自動採点 [8] など多岐に渡る．これらの研究は，作文の評価時における負荷を軽減することを目的としており，この点で本研究の目的とは異なる．

ゲーミフィケーション・GWAP については，これに関するワークショップが LREC2020 で開かれるなど，様々な目的で取り組まれている．例えば，常識獲得 [9, 10] や語義曖昧性の解消 [11] などがある．ことばつなぎゲームは，アノテーションデータの収集より，能力の育成に焦点を当てている．

5　おわりに

本研究では，大規模な言語資源を利用した作文のゲーミフィケーションに取り組んだ．小学生の作文教育に向けたゲームを考案し，そのためのデータおよびデモシステムを構築した．デモシステムをもとにアプリを作る予定であり，今後，実用性を検証していきたいと考えている．また，このゲームを通した未知の関係文の収集も検討する．

謝辞　本研究は，(公財) 日本漢字能力検定協会の支援を受けています．

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

参考文献

[1] 立命館大学図書館. レファレンス事例詳細, 2017. https://crd.ndl.go.jp/reference/modules/d3ndlcrdentry/index.php?page=ref_view&id=1000209543.

[2] 水谷勇介, 河原大輔, 黒橋禎夫. クラウドソーシングを用いた習得時期の想起質問に基づく単語難易度データベースの構築. 言語処理学会 第 25 回年次大会, 2019.

[3] Kazumasa Omura, Daisuke Kawahara, and Sadao Kurohashi. A Method for Building a Commonsense Inference Dataset based on Basic Events. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020.

[4] Daisuke Kawahara, Daniel Peterson, Octavian Popescu, and Martha Palmer. Inducing Example-based Semantic Frames from a Massive Amount of Verb Uses. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, 2014.

[5] 村山太一, 若宮翔子, 荒牧英治. WORD GINI: 語の使用の偏りを捉える指標の提案とその応用. 言語処理学会 第 24 回年次大会, 2018.

[6] Arseny Tolmachev, Daisuke Kawahara, and Sadao Kurohashi. Design and Structure of The Juman++ Morphological Analyzer. *Journal of Natural Language Processing*, Vol. 27, No. 1, pp. 89–132, 2019.

[7] 小川耀一朗, 山本和英. 日本語文法誤り訂正における誤り傾向を考慮した擬似誤り生成. 言語処理学会 第 26 回年次大会, 2020.

[8] Tomoya Mizumoto, Hiroki Ouchi, Yoriko Isobe, Paul Reisert, Ryo Nagata, Satoshi Sekine, and Kentaro Inui. Analytic Score Prediction and Justification Identification in Automated Short Answer Scoring. In *Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications*, 2019.

[9] Yen-ling Kuo, Jong-Chuan Lee, Kai-yang Chiang, Rex Wang, Edward Shen, Cheng-wei Chan, and Jane Yung-jen Hsu. Community-Based Game Design: Experiments on Social Games for Commonsense Data Collection. In *Proceedings of the ACM SIGKDD Workshop on Human Computation*, 2009.

[10] Naoki Otani, Daisuke Kawahara, Sadao Kurohashi, Nobuhiro Kaji, and Manabu Sassano. Large-Scale Acquisition of Commonsense Knowledge via a Quiz Game on a Dialogue System. In *Proceedings of the Open Knowledge Base and Question Answering Workshop (OKBQA 2016)*, 2016.

[11] Noortje J. Venhuizen, Valerio Basile, Kilian Evang, and Johan Bos. Gamification for Word Sense Labeling. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013) – Short Papers*, 2013.

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).