

# 名詞と動詞の組み合わせに基づく 「語彙組み合わせ問題」の自動作問 Automatic Generation of Combinational Items Based on Combinations of Japanese Nouns and Verbs

張 寓杰<sup>1</sup>, 加藤 健太郎<sup>1</sup>, 寺井 あすか<sup>2</sup>, 中川 正宣<sup>3</sup>  
Yujie Zhang, Kentaro Kato, Asuka Terai, Masanori Nakagawa

<sup>1</sup>ベネッセ教育総合研究所, <sup>2</sup>公立ほこだて未来大学システム情報科学部,  
<sup>3</sup>大妻女子大学人間生活文化研究所

Benesse Educational Research and Development Institute, Future University Hakodate, Otsuma Women's University  
zhangyj@mail.benesse.co.jp

## Abstract

The present study proposes an automatic item generation system which creates the two types of combinational items, the subject(noun)-predicate(verb) combinations and the object(noun)-verb combinations, based on the co-occurrence frequencies between Japanese nouns and verbs. Those co-occurrence frequencies have been computed by the statistical analysis of large scale Japanese language data from a previous study.

The present system, first selects proper verbs which have high frequency with the initial input nouns to generate the correct combinational items. Then, the system selects nouns which have high frequency with the selected verbs to generate the other correct combinational items. In the next step, the system selects the nouns which have zero co-occurrence frequency with the selected verbs to generate the wrong combinational items. In those procedures, the difficulties of the correct items are controlled by the abstractness of the initial and selected nouns, while the difficulties of the wrong items are controlled by the similarities between nouns of the correct items and those of the wrong items.

In a practical test consisting of the two types of generated combinational items, 54 university students were asked to answer whether each item was correct or wrong. The results show that the difficulties controlled in the system are sufficiently consistent with the correct answer ratios of the students.

**Keywords** — Automatic Generation, Combinational Item

## 1. 研究背景と研究目的

近年, ICT 分野の急速な進歩に伴い, 学校現場でも学習支援システムや, CBT 試験(Computer-Based Testing)が導入され, より多くの問題を迅速かつ大量に生成することが求められている. 英単語の意味に関する自動作問領域ではすでに多くの研究[1,2,3,5,9]が発表されているが, 日本語に関する自動作問の研究はまだ多くはない. たとえば, 久田・前田 [7]は, ニュース記事を用いた時事問題の自動作問システムを提案している. また, 林・垂水[8]では学習者が多肢選択問題を

作成, 投稿して, それらの問題を学習者間で共有できる自主学習素材共有システムを開発している. しかし, 日本語の語彙そのものに関する自動作問の研究はほとんど行われていないのが現状である.

日本語の語彙テストとしては既に, 朝日新聞社とベネッセコーポレーションが共同で実施している「語彙・読解力検定[6]」がある. この「語彙・読解力検定」では, 一単語の意味を問う問題や, 文章中の単語の意味を問う問題を出题している. 一方, 清水・萩原 [11]では, 言葉の印象の理解が, 人間の感情理解や比喩理解において大きな役割を果たすことが指摘されている. 特に動詞は名詞と共起することで, 言葉の印象を大きく変えることがある. そこで, この研究では名詞と動詞の組み合わせを正しく理解できれば, 言葉の深層に含まれる意図や事実を推測することができると仮定し, 名詞と動詞の組み合わせに対する印象の自動推定システムを提案している.

確かに, 上記の研究で指摘されているように, 単語の意味は他の単語との組み合わせによって大きく変わってしまうことが多い. 特にこの変化は, 名詞と動詞の組み合わせにおいて顕著である. たとえば「走る」という極めて基本的な意味の動詞でも, 「車が走る」という表現の場合と「緊張が走る」という表現の場合では, その意味が異なっている. 一方, 「緊迫が走る」という表現は, 字面は似ていても適切な表現とは言えない. これらの点を考慮すると, 同じ「走る」という動詞がどんな名詞と組み合わせると適切な表現になるのかという組み合わせの問題が極めて重要であると言える.

つまり, 個々の単語の基本的な意味を知っているだけでは言語の理解能力としては不十分である. 少なくともどの名詞とどの動詞の組み合わせが言語として意味を成すかを知っていることは, 言語を正しく話し書くために必要不可欠な知識であると言える.

本研究では、特に名詞と動詞の組み合わせに着目し、それらの組み合わせの適切さが理解されているかどうかを測定するための「語彙組み合わせ問題」の作成を試みる。本研究における「語彙組み合わせ問題」とは、例えば「緊張が広がる」のような名詞と動詞の組み合わせ語句が日本語として適切かどうかを答える問題形式を意味する（以下では適切な組み合わせを正解、不適切な組み合わせを誤りとする）。

しかし、新聞や文学作品等で一般的に用いられる名詞や動詞は各々数万を超えている。さらに、「主語と述語」、「目的語と動詞」の場合では各々その組み合わせパターンは異なっている。単に辞書を参照するだけで、名詞と動詞の適切な組み合わせを組織的に構成し、組み合わせ問題を作成することは極めて困難である。さらに、問題作成では単なる大量生成ではなく、難易度が異なる問題の生成が求められている[10]。

本研究では、上記の問題を解決するため、新聞データ（毎日新聞 1991年-2008年）や文学作品を含む大規模言語コーパスから抽出した、三万以上の名詞と七万に及ぶ動詞(表 1 参照)の組み合わせの共起頻度のデータベースを用いて、名詞と動詞の「語彙組み合わせ問題」を難易度別に自動的に作成できるシステムを提案する。

表 1 本研究で用いた共起頻度データに含まれる異なる名詞と動詞の語数

	名詞（目的語）-を- 動詞（述語）	名詞（主語）-が- 動詞（述語）
名詞数	38,816	34,668
動詞数	92,567	70,398

本研究の前提として、既に、上記の日本語大規模コーパスの言語統計解析に基づく、名詞、動詞、形容詞に関する共起頻度データベースおよび、確率的言語知識構造(潜在クラス群)が作成されている[4]。本研究で提案するシステムでは、この共起頻度のデータベースと潜在クラス群に基づく以下のような一連の手続きに従い、正誤の語彙組み合わせ問題を難易度別に自動的に作成する。

## 2. 研究方法

前記の共起頻度データにおける日本語の語彙組み合わせは「名詞（主語）が動詞（述語）」、「名詞（目的語）を動詞」、「名詞（目的語）に動詞」三種類があるが、本

研究は自動作問の最初の試みとして、「名詞（主語）が動詞（述語）」と「名詞（目的語）を動詞」二種類の組み合わせについて、以下に示す手順で自動作問を行う。

### 2.1 抽象的な名詞から始める組み合わせの自動作問

この方法では、まず、抽象的な名詞を初期値として入力することにより、同じ動詞を用いて、「抽象的な意味を持つ語彙組み合わせ」と「具体的な意味を持つ語彙組み合わせ」の二種類の組み合わせを作ることができる。

今回は、まず「抽象的な意味を持つ語彙組み合わせ」として4個の抽象的な名詞各々に4個の動詞を組み合わせ、16個の組み合わせを自動的に生成する。次に「具体的な意味を持つ語彙組み合わせ」として、同じ4個の動詞各々に4個の具体的な名詞を組み合わせ、16個の組み合わせを生成する。すなわち、「名詞（主語）が動詞（述語）」、「名詞（目的語）を動詞」各々で32個の組み合わせが自動的に生成されることになる。

#### 2.1.1 「名詞（主語）が動詞（述語）」の場合の自動作問の手順

「名詞（主語）が動詞（述語）」の場合の自動作問の手順を図1に示す。

①まず、確率的言語構造を構成する潜在クラスのなかから、比較的抽象的な名詞群の所属確率が高い潜在クラスを複数個抽出する。

②次に、上記の潜在クラスの一つをランダムに選定し、そのクラスへの所属確率が高い名詞を任意個数抽出する。抽象度が高い名詞を用いると、難易度が高い組み合わせを作ることができると期待される。

例1:「気運、機運、関心、緊張、緊迫、内圧」など

③次に、これらの名詞の各々を主語として「が」で繋がる共起頻度が高い動詞群を選出する。

例2:「走る、生まれる、盛り上がる、集まる、高まる」など

④例1の名詞と共起頻度が高い動詞(例2の動詞)を組み合わせると、以下のような比較的難易度の高い問題が作成できる。

例3:「緊張が走る」、「機運が生まれる」、「気運が盛り上がる」、「関心が集まる」、「内圧が高まる」

⑤次に、例2の動詞群の中から、特に共起頻度の高い動詞群を選別し、それらの動詞の主語として「が」で繋がる共起頻度が高い名詞群を選出する。

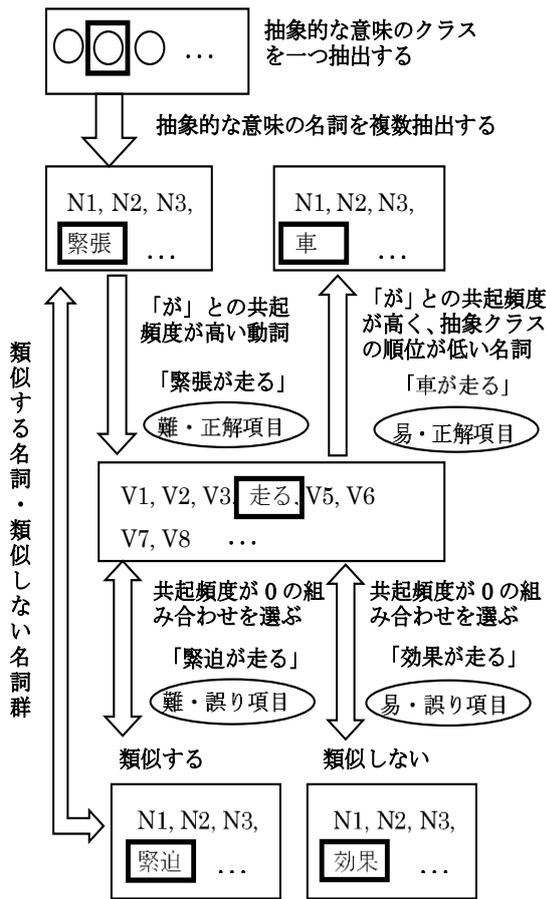


図1 抽象的な名詞から始める組み合わせの自動作問手順：「名詞(主語)が動詞(述語)」の場合

⑥次は、⑤で選んだ名詞群から、比較的、抽象的な潜在クラスでの順位が低い名詞（「車、子供、気持、人、気分」等）を選出し、上記の動詞と組み合わせる。同じ動詞に対して抽象度が低い名詞を組み合わせると、より難易度の低い組み合わせを作ることができると期待される。

例4：「車が走る」、「子供が生まれる」、「気持ちが盛り上がる」、「人が集まる」、「気分が高まる」

⑦さらに、②で選定したのと同じ潜在クラスに属する名詞群と③で選出した動詞群での共起頻度が0の組み合わせを選び、誤り問題を作成する（例えば「緊迫が走る」）。この誤り問題は、正しい組み合わせ（例えば「緊張が走る」）と名詞が類似していて、かつ同じ動詞（「走る」）を用いている。そのため正誤の判断がそれだけ難しいと考えられる。

### 2.1.2 「名詞（目的語）を動詞」組み合わせの自動作問の手順

図2のように、「名詞（主語）が動詞（述語）」の場

合と同様の手続きで自動的に問題を作成する。

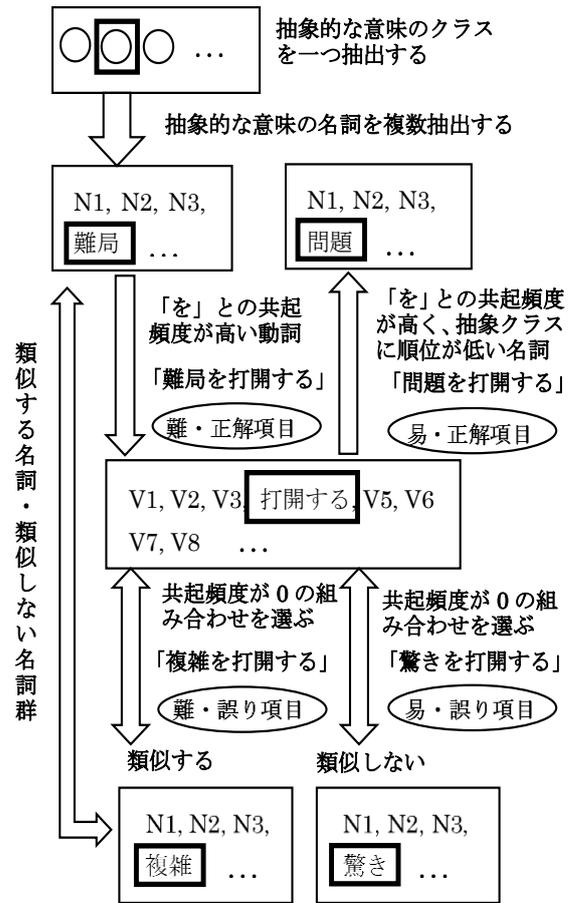


図2 抽象的な名詞から始める組み合わせの自動作問手順：「名詞(目的語)を動詞」の場合

## 2.2 具体的な名詞から始める組み合わせの自動作問

この方法では、2.1の方法とは逆に、具体的な名詞の所属確率が高い潜在クラスに属する任意個の名詞群を初期値として用いて問題を作成した。この方法では、まず「名詞（主語）が動詞（述語）」、「名詞（目的語）を動詞」それぞれ、4個の名詞、各々に4個の動詞を組み合わせ、16個の組み合わせを自動的に生成する。さらにそれぞれ2個の組み合わせを追加し2.1の場合と合わせて各々計50個の組み合わせが生成される。

この方法における「名詞（主語）が動詞（述語）」の自動作問の手続きは図3に示すとおりである。

①まず、確率的言語構造で「名詞(主語)が動詞(述語)」に対応した名詞の潜在クラスの中から比較的具体的な名詞群の所属確率が高い潜在クラスの一つを選び、そのクラスへの所属確率が高い名詞を任意個数選出する。

例1：「大統領、議長、首相、書記」など

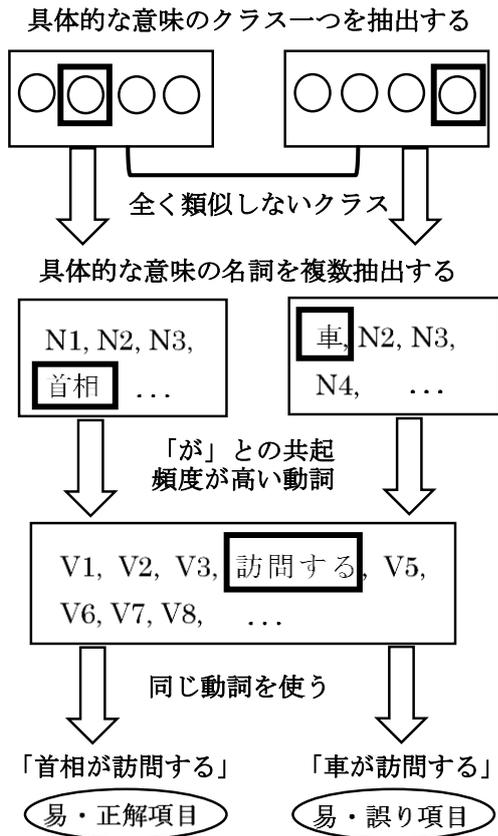


図3 具体的な名詞から始める組み合わせの自動作問手順：「名詞(主語)が動詞(述語)」の場合

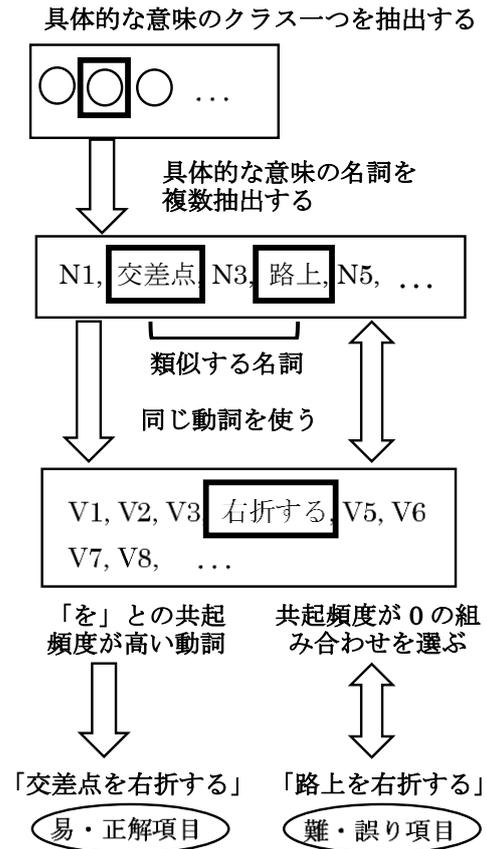


図4 具体的な名詞から始める組み合わせの自動作問手順：「名詞(目的語)を動詞」の場合

②次に、これらの名詞の各々を主語として「が」で繋がる共起頻度が高い動詞群を選出する。

例2：「訪問する，発表する，表明する」など

③例1の名詞と共起頻度が高い動詞(例2の動詞)を組み合わせると、以下のような比較的難易度の低い問題が作成できる。

例3：「大統領が訪問する」，「議長が発表する」，「首相が表明する」など

④さらに、①で選定した潜在クラスとはまったく類似しない別の具体的な意味を持つ潜在クラスに属する名詞群の中から、②で選出した動詞群との共起頻度が0の組み合わせを選び、誤り問題を作成する(例えば「ダンプカーが訪問する」)。この誤り問題は、正しい組み合わせ(例えば「大統領が訪問する」)とは名詞がまったく異なっているが、あえて同じ動詞(「訪問する」)を用いている。そのため正誤の判断がそれだけ易しいと考えられる。

一方、「名詞(目的語)を動詞」の自動作問手順は図4に示すとおりである。

①まず、確率的言語構造で「名詞(目的語)を動詞」に対応した名詞の潜在クラスの中から比較的具体的名詞群の所属確率が高い潜在クラスの一つを選び、そのクラスへの所属確率が高い名詞を任意個数選出する。

例1：「水中，海上，歩道，交差点」など

②次に、これらの名詞の各々を目的語として「を」で繋がる共起頻度が高い動詞群を選出する。

例2：「泳ぐ，北上する，歩く，右折する」など

③①で選んだ名詞と共起頻度が高い動詞(例2の動詞)を組み合わせると、以下のような比較的難易度の低い問題が作成できる。

例3：「水中を泳ぐ」，「海上を航行する」，「廊下を歩く」，「交差点を右折する」など

④さらに、①で選定した具体性の高い潜在クラスと同じ潜在クラスに属する名詞群の中から②で選出した動詞群との共起頻度が0の組み合わせを選び、誤り問題を作成する(例えば「廊下を右折する」)。この誤り問題は、正しい組み合わせ(例えば「交差点を右折する」)と名詞が類似していて、かつ同じ動詞(「右折する」)を用いている。そのため正誤の判断がそれだけ難しいと

考えられる。

### 3. 実験方法

以上の手順を用いて自動生成した問題に関して、システムが想定した難易度が、実際の人間の判断と一致するかどうかを検証するため、大学一年生 54 名を実験参加者として Web 上で実験調査を実施した。

実験材料としては上記 2.1 と 2.2 の手続きに従って生成された、難易度が異なる組み合わせ項目を用いた。「名詞(主語)が動詞(述語)」、「名詞(目的語)を動詞」各々につき 50 項目を生成し、全部で 100 項目の実験課題を作成した。実験参加者は各項目に対して、日本語として「正しい」、「正しくない」、「分からない」、3つの選択肢から一つを選択した。また、以降の結果の集計では、正解項目を正しい、誤り項目を正しくないと判断した場合を正答とし、正解項目、誤り項目共に「分からない」を選択した場合はすべて正答に含めず、正答率を計算した。

## 4. 結果と考察

### 4.1 抽象的な名詞から始める組み合わせの実験結果

#### 4.1.1 「名詞(主語)が動詞(述語)」の場合

2.1.1 で作成した抽象的名詞群、具体的名詞群の正解項目、誤り項目それぞれの正答率の平均値は表 2 に示すとおりである。

この実験結果の正答率の逆正弦変化後の値に対して、抽象・具体、正解項目・誤り項目の二元配置の分散分析を行った結果、抽象・具体要因に有意水準 1 パーセントで有意差があり ( $F(1, 29)=10.23$ ,  $p=0.003$ ,  $\eta^2=0.25$ )、正解項目・誤り項目要因 ( $F(1, 29)=0.184$ ,  $p=0.19$ ,  $\eta^2=0.04$ ) 及び交互作用 ( $F(1, 29)=0.03$ ,  $p=0.87$ ,  $\eta^2=0.0006$ ) には有意差がなかった。

表 2 「名詞(主語)が動詞(述語)」の正答率の平均値

	正解項目		誤り項目	
	N	Mean	N	Mean
抽象的名詞群	7	0.642	9	0.537
具体的名詞群	9	0.877	8	0.785

上記の表 2 及び分散分析の結果から、抽象的名詞群の方が具体的名詞群より正解項目・誤り項目共に正答

率が低い。つまり抽象的名詞群の方が、表現として正しいかどうかを正しく判断できないという意味では、難易度が高いといえる。たとえば、抽象的名詞との組み合わせ、「気運が盛り上がる」という正解項目の正答率は 0.2 と非常に低く、一方、同じ動詞と具体的名詞との組み合わせ正解項目「議論が盛り上がる」の正答率は 0.93 と非常に高い。

### 4.1.2 「名詞(目的語)を動詞」の場合

2.1.2. で作成した抽象的名詞群、具体的名詞群の正解項目、誤り項目それぞれの正答率の平均値は表 3 に示すとおりである。

表 3 「名詞(目的語)を動詞」の正答率の平均値

	正解項目		誤り項目	
	N	Mean	N	Mean
抽象的名詞群	5	0.485	11	0.620
具体的名詞群	9	0.689	8	0.831

この実験結果の正答率の逆正弦変化後の値に対して、やはり、抽象・具体、正解項目・誤り項目の二元配置の分散分析を行った結果、抽象・具体、正解項目・誤り項目の二元配置の分散分析の結果では、抽象・具体要因に有意水準 1 パーセントで有意差があり ( $F(1, 29)=7.83$ ,  $p=0.009$ ,  $\eta^2=0.20$ )、正解項目・誤り項目要因 ( $F(1, 29)=2.98$ ,  $p=0.095$ ,  $\eta^2=0.075$ )、及び交互作用 ( $F(1, 29)=0.017$ ,  $p=0.90$ ,  $\eta^2=0.0004$ ) には有意差がなかった。

上記の表 3 及び分散分析の結果から、抽象的名詞群の方が具体的名詞群より正答課題、誤答課題共に正答率が低い、つまり難易度が高いことがわかる。具体的には、抽象的名詞との組み合わせ正解項目「熾烈をきわめる」の正答率は 0.35 で非常に低く、反対に同じ動詞と具体的名詞との組み合わせ正解項目「道をきわめる」の正答率は 0.7 でかなり高い。

### 4.2 具体的な名詞から始める組み合わせの実験結果

#### 4.2.1 「名詞(主語)が動詞(述語)」の場合

2.2.1. で作成した具体的名詞群(単独)の正解項目、誤り項目それぞれの正答率の平均値は表 4 に示すとおりである。

この場合の正答率の逆正弦変化値と、4.1.1の正答率の逆正弦変化値とを合わせて、抽象・具体・具体単独、正解項目・誤り項目の二元配置の分散分析を行った結果、抽象・具体・具体単独要因に有意水準0.1パーセントで有意差があり ( $F(2, 44)=16.35, p=0.00, \eta^2=0.41$ ), 正解項目・誤り項目要因 ( $F(1, 44)=0.61, p=0.44, \eta^2=0.008$ ), 及び交互作用 ( $F(2, 44)=1.52, p=0.23, \eta^2=0.04$ ) には有意差がなかった。また、下位検定の結果、抽象的名詞群と具体的名詞群の間 ( $t(44)=5.62, p=0.00$ ), 及び抽象的名詞群と具体的名詞単独群の間 ( $t(44)=3.82, p=0.00$ ), にそれぞれ有意水準0.1パーセントで有意差があった。一方、具体的名詞群と具体的名詞単独群の間には有意な差はなかった。

表4 「名詞(主語)が動詞(述語)」  
(具体的名詞群(単独))の正答率の平均値

	正解項目		誤り項目	
	N	Mean	N	Mean
具体的名詞群 (単独)	9	0.930	8	0.968

上記の表2, 表4及び分散分析の結果から、やはり抽象的名詞群の方が具体的名詞群より正解項目、誤り項目共に正答率が低い、つまり難易度が高いことがわかる。さらに、具体的名詞群(単独)より、抽象的名詞群の方は難易度が高いこともわかる。特に具体的名詞群(単独)では、正解項目と誤り項目で、同じ動詞に対してまったく異なる種類の名詞を用いているため、誤りの項目の正答率が高くなっていると考えられる。例えば、正解項目「大統領が訪問する」の正答率は0.93でかなり高いが、同じ動詞に対応する誤り項目「列車が訪問する」の正答率は0.96とさらに高くなっている。

#### 4.2.2 「名詞(目的語)を動詞」の場合

2.2.2. で作成した具体的名詞群(単独)の正解項目、誤り項目それぞれの正答率の平均値は表5に示すとおりである。

この場合の正答率の逆正弦変化値と、4.1.2の正答率の逆正弦変化値とを合わせて、抽象・具体・具体単独、正解項目・誤り項目の二元配置の分散分析を行った結果、抽象・具体・具体単独要因 ( $F(2, 44)=4.21, p=0.02, \eta^2=0.14$ ) と交互作用 ( $F(2, 44)=4.18, p=0.02, \eta^2=0.14$ ) に有意水準5パーセントで有意差があり、正解項目・

誤り項目要因には有意差がなかった ( $F(1, 44)=0.02, p=0.89, \eta^2=0.0003$ )。

表5 「名詞(目的語)を動詞」  
(具体的名詞群(単独))の正答率の平均値

	正解項目		誤り項目	
	N	Mean	N	Mean
具体的名詞群 (単独)	5	0.904	12	0.611

また、下位検定の結果、抽象的名詞群と具体的名詞群の間 ( $t(44)=2.57, p=0.014$ ), 及び抽象的名詞群と具体的名詞単独群の間 ( $t(44)=2.50, p=0.016$ ) にそれぞれ有意水準5パーセントで有意差があった。特に、交互作用に関する下位検定では、具体的名詞単独群において正答課題と誤答課題の間に、有意水準5パーセントで有意差があることがわかった ( $F(1, 44)=5.76, p=0.02, \eta^2=0.095$ )。上記の表3, 表5及び分散分析の結果から、やはり抽象的名詞群の方が具体的名詞群より正答課題、誤答課題共に正答率が低い、つまり難易度が高いことがわかる。さらに、交互作用を考慮した下位検定の結果から、特に具体的名詞群(単独)では、誤答課題の方が正答課題より難易度が高いことがわかる。

具体的名詞群単独の場合のこの正答率の差は、誤り項目の名詞が同じ動詞を組み合わせた正解項目の名詞と類似している場合に顕著である。たとえば、「交差点を右折する」という正解項目の正答率は0.94で非常に高いが、「交差点」と類似した名詞と同じ動詞とを組み合わせた誤り項目「歩道を右折する」の正答率は0.3と非常に低い。同じく、正解項目「海上を北上する」の正答率は0.78でかなり高いのに、類似した名詞を用いた誤り項目「路上を北上する」の正答率は0.33でかなり低いと言える。

以上の結果から、本研究の自動作問のシステムを用いて生成された「名詞(主語)が動詞(述語)」と「名詞(目的語)を動詞」二種類の組み合わせ問題について、システムで想定された難易度が実際の人間にとっての難易度と十分一致していることが、実験に基づき実証された。すなわち、最初に入力した抽象的な名詞群と動詞の組み合わせ問題は、同じ動詞に対応する具体的名詞群との組み合わせよりも難易度が高い。また、具体的名詞群を用いる場合でも、誤り項目の名詞が同じ動詞に対応する正解項目の名詞と類似している場合は、類

似していない場合より、難易度が高くなることも確認された。これらの結果から、今回開発した自動作問システムが難易度の自動設定についても十分有効に機能していることが実証されたわけで、今後の当該システムの教育現場での語学テスト問題の自動作成等に向けた実用化の可能性が高いと言える。

## 5. 今後の課題

本研究では語彙組み合わせの難易度付き自動作問の最初の試みとして、「名詞(主語)が動詞(述語)」と「名詞(目的語)を動詞」二種類の組み合わせについて、同じ手順で自動作問を行った。今後は「名詞(目的語)に動詞」あるいは「形容詞-名詞」の組み合わせでも同様の手順で自動作問を行うことができる。

また、本システムの特徴として、初期入力で用いる名詞群のレベルを変えることで、様々な難易度の組み合わせの項目を生成することができる。例えば、小学校、中学校、高校それぞれにおいて基本的に学習する用語を初期入力し、それぞれの段階に応じた難易度の組み合わせ項目を作成することも可能である。また外国人の日本語能力の段階に応じた名詞を入力し、段階的な難易度の項目生成にも対応できる。

さらに、本研究で構成した組み合わせ項目の自動生成の方法は主に、名詞と動詞の文法関係に基づく共起頻度を用いている。すなわち、名詞と動詞の文法関係に基づく共起頻度に関する十分に大量のデータが取得できれば、この方法はどんな言語にも応用可能である。

ただし、言語の意味は、時代や世代、環境等さまざまな社会的要因によって大きく影響を受ける。このような社会的要因に由来する言語の意味の変化を考慮に入れて、今後、共起頻度を計算するための言語データベースとして、インターネット等の新たな資源からの言語データの取得も考えている。

また、本研究で使用した実験材料は100問かつ被験者は54人で、項目数と被験者数ともに十分とは言えない。今後は項目数と被験者を増やして検証する必要がある。

今後は新たに加えた語彙に関する共起頻度データに基づき、本研究で提案した自動作問の方法を用いて、初期設定で入力する名詞群を様々に変えることで、課題内容、難易共に多様なテスト項目の自動生成とその実証調査を行う予定である。

さらに、将来、本方法を日本語以外、中国語や英語等様々な言語へ適用し、多言語での組み合わせ項目の自

動生成システムの構成も試みていきたい。

## 参考文献

- [1] Agarwal, M. and Mannem, P. (2011). Automatic gap-fill question generation from text books. In Proceedings of the Sixth Workshop on Innovative Use of NLP for Building Educational Applications, pages 56–64.
- [2] Brown, J. C., Frishkoff, G. A., and Eskenazi, M. (2005). Automatic question generation for vocabulary assessment. In Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, pages 819–826.
- [3] Chen, W., Aist, G., and Mostow, J. (2009). Generating questions automatically from information text. In Proceedings of AIED 2009 Workshop on Question Generation, pages 17–24.
- [4] 張寓杰, 寺井あすか, 董媛, 王月, 中川正宣. (2013). 日本語と中国語における帰納的推論の比較研究—言語統計解析に基づく計算モデルを用いて—. 認知科学, No.20, Vol.4, 439-469.
- [5] Gierl, M. J., & Haladyna, T.M. (2013). Automatic item generation. New York, NY: Routledge.
- [6] 語彙・読解力検定. <http://www.goi-dokkai.jp/>
- [7] 久田将史, 前田亮. (2007). ニュース記事を用いた時事問題自動作問システム. DEIM Forum A2-1.
- [8] 林敏浩, 垂水浩幸. (2012). 自主学習素材共有システムにおける自動作問機能. 教育システム情報学会第37回全国大会, 348-349.
- [9] Holling, H., Bertling, J. P., & Zeuch, N. (2009). Automatic item generation of probability word problems. Studies in Educational Evaluation, 35, 71–76.
- [10] Kane, M.T. (2006). Validation. In R. L. Brennan(Ed.), Educational measurement (4th ed., pp.17-64). Westport, CT: Praeger Publishers.
- [11] 清水浩平, 萩原将文. (2011). 名詞と動詞の組み合わせに対する印象推定法. 日本感性工学会論文誌, Vol.10, No.4, pp.505-514.