

令和元年6月18日現在

機関番号：32689

研究種目：若手研究(B)

研究期間：2015～2018

課題番号：15K16786

研究課題名(和文) 留学生のための、特殊コーパスを用いた、基本的な専門語の新規抽出法とその活用

研究課題名(英文) Method and practice on extracting subject-specific terms from domain-specific corpora for international student education

研究代表者

喜古 正士 (KIKO, Masashi)

早稲田大学・総合研究機構・その他(招聘研究員)

研究者番号：90632183

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：これまでの専門語研究においては、専門語が一般的に名詞句相等であるという事情から、修飾語句はあまり着目されてこなかった。しかし、本研究を通して分析対象テキストを適切に選ぶことで、それらの修飾語句を効果的に抽出できることが確認された。また、一般語が専門的な意味を担っているもの(本研究では“文脈依存する専門語”と呼ぶ)も多くそのような形で確認された。加えて、テキストのジャンルとして専門分野の異なりが文体に影響していることを、品詞構成比率の異なりから間接的に示した。これらの調査を踏まえ、母語話者・非母語話者の別を超えて、“教授言語としての日本語”という枠組みを考える必要性を提起した。

研究成果の学術的意義や社会的意義

専門語は一般的に名詞句相等であるという事情から、これまでの専門語研究においては修飾語句はあまり着目されてこなかった。これに対して、分析対象テキストを適切に選択することで効果的に議論できることを示した。また、一般語が専門的な意味を担っているものも多くそのような形で確認され、専門語研究の新しい方法を提案できた。

また実用的には、日本語を母語としない人を日本国内で活躍する高度人材として受け入れるに際し、「日本語による専門的なコミュニケーション」に必要な語彙を提案するための基盤作りに活用できる。

研究成果の概要(英文)： Since technical terms are usually noun phrases, we rarely take modifiers into consideration. The results of this research show that we can derive the subject-specific vocabulary as a whole including such special modifiers, by making appropriate domain-specific corpora. It also contain general words which play special role in the context: “context-dependent terms.” In addition, part-of-speech ratios in the corpora indirectly show that the difference in subjects make the difference in styles.

Taking these differences into account, “Japanese as a language of instruction” is emerged as a key concept. We should explore it for native and non-native speakers.

研究分野：日本語教育

キーワード：日本語教育 専門語 専門日本語

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

物理の文脈で「振り子を静かに手放した」とあれば、「静かに」という語は速度が0であることを示唆する基本的な語である。しかしながら、学習用の用語集や専門語辞典でもこの語積が載ることは希である。当然、国語辞書にはこのような用法の記述は見られない。一般語が文脈上の専門的な意味を担うこのような用法は、多くの場合、使用場面を通して学ばれる。このように、文脈上の専門的な意味を担っている一般語をここでは“文脈依存する専門語”と呼ぶこととする。

外国人留学生のように、日本語を母語としない人が「日本語による専門的なコミュニケーション」を求められる際、その専門家集団が用いる語彙・文法の集まりとしての“専門語”をどれだけ理解しているかが重要になる。そこで、“文脈依存する専門語”を含めた“専門語”の総体を明らかにすることが「日本語で学ぶ」手助けとして重要になってくる。

2. 研究の目的

留学生の教育を考えると、学ぶべき“基本的な専門語”を提示することは、学習効率の点からも重要である。しかしながら、先に提示した“文脈依存する専門語”は、辞書や索引の項目としてあまり取り上げられてこなかったという経緯もあり、これまでの専門語研究では見落とされてきた。では、広範に用いられるにも関わらず辞書に載らない語の用法は、どの程度存在し、またどのように獲得されているのだろうか。そして、そのような語はどこに、またなぜ偏在しているのか。このような素朴な疑問から、ある種の「専門的なコミュニケーション」をしていると話者同士が了解しているとき、話題を限定・共有するために“専門語”が多く用いられるであろうという仮説が立てられた。

そこで本研究では、記述的な「教科書」よりも、より双方向性を有する「試験問題」の方がより「専門的なコミュニケーション」が成立すると考え、その範疇での“専門語”の振る舞いを考察するものとした。具体的には、理工系留学生のための、“文脈依存する専門語”も含めた“基本的な専門語”の新規抽出方法とその効果を検証することを目的とし、対象を「大学入学程度の物理学の語彙」に絞り調査する。

3. 研究の方法

(1) “文脈依存する専門語”の特徴を捉える

専門語の調査には特殊コーパスを用いることが一般的である。中等教育程度の語彙を調査しようとするとき、国立国語研究所の統計資料(高校教科書の語彙調査)や検定教科書の本文や索引を利用するなどの工夫が必要となる。しかし、これらのコーパス類から“文脈依存する専門語”を調査しようすると、講義中しばしば用いられるにも関わらず、出現頻度が著しく低いという問題に遭遇する。これは、“文脈依存する専門語”が話題を限定・共有する際に多く用いられるという性格から、記述された教科書本文よりも、より双方向的な講義や問題の設問で頻度が上昇することで説明される。そこで、本研究では

・日本留学試験問題文の調査

を行うことで、“基本的な専門語”の抽出効率が上昇することを確認する。具体的には、出版されている『日本留学試験』(2011~2014)の「物理」部分の問題文を用いてコーパスを構築し、語彙の調査を行った。

(2) 使用場面による、使用語彙の変化を捉える

“文脈依存する専門語”の出現頻度が問題の設問で上昇することを上で述べたが、そもそもテキストの種類が変わることで、同じ分野でも使用語彙の構成比率がどう変化しているかを調べるため、

・教科書、問題集、試験問題の比較調査

を行った。具体的には、検定教科書『改訂版高等学校物理Ⅰ,Ⅱ』(数研出版)の本文、傍用問題集『三訂版リード 物理Ⅰ・Ⅱ』(数研出版)の問題文、センター試験『大学入試センター試験本試験(物理)』(2007~2013)の問題文を用いてコーパスを構築し、語彙の調査を行った。

(3) 「物理の日本語」における、使用語彙の特徴を捉える

本研究では調査の対象を「大学入学程度の物理学の語彙」としているが、「物理の日本語」の特徴を他の科目との比較から検討するために、

・センター試験問題文の科目間比較調査

を行った。具体的には、『大学入試センター試験本試験』(2007~2013)の科目毎(国語は現代

文を対象とし、評論と小説を別にした)の問題文を用いてコーパスを構築し、品詞構成比率の調査を行った。

(4)物理翻訳語の特徴を考える

一般語の専門的文脈での借用が初期の翻訳段階でどの程度起きたのかを調べるため、

・明治期翻訳資料の調査

を行った。具体的には、『物理学術語和英仏独対訳字書』の調査報告と現物とを付き合わせながら、物理翻訳語に用いられた構成語基の確認を行い、データベースを作成した。

4. 研究成果

(1) “文脈依存する専門語”

まず、作成した問題集・試験問題のコーパスから、教科書本文に比べ、“専門語”の比率は問題文で向上することが確認できる。また、“文脈依存する専門語”が話題を限定・共有する際に多く用いられるという性格から、「相の類」に限った集計を取ると、テキスト種別により出現頻度が大きく変化することが見て取れる(表1)。

また、物理翻訳語の成立過程で積極的に既存語を転用して当てた可能性が確認されたが、“文脈依存する専門語”の多くが「修飾語句」の形で見られることの直接の説明には至っていない。漢語語基の造語力を活用したため、形容動詞語幹の意味の多重化が起こった可能性も含めてさらなる検証が必要である。

(2)媒体毎の特徴

図1は各テキスト(検定教科書・傍用問題集・日本留学試験(物理)・大学入試センター試験本試験(物理))をコーパスとして、それぞれの「異なり語」による品詞構成比率を表したものである。自立語を体の類・用の類・相の類・その他の4類に分け、横軸に「体の類」の百分率を取り、「用の類」・「相の類」の百分率を縦軸に取った散布図となっている。図1の特徴として散布図の点が一直線上に並んで見える。これは、体・用・相のどれか一つの比率からテキスト全体の品詞構成が決まることを示しており、“大野・水谷の法則”と一致する。

また、品詞構成比率の異なりがある程度文体の差異を反映していると考えられるなら、「問題文は出題者と受験者とが話題を限定・共有する構造にあり、修飾構造が多く複雑である」という仮説をサポートするものとなっている。センター試験の問題文の構造が難しいことも現れていると考えられる。

加えて、国立国語研究所による『現代日本語書き言葉均衡コーパス語彙表』(BCCWJ語彙表)の「図書館・書籍レジスター」を対照コーパスとして、「センター試験」・「留学試験」・「問題集」において、有意水準0.5%の対数尤度比検定で特徴語を決定した。特徴語のうち相の類のもの(異なりで73語)で、3つのテキストに共通する27語を、使われ方に応じて3クラスに分類し表2に掲載した。加えて、「教科書」を対照コーパスとして同様の検定を行った(表3)ところ、“文脈依存する専門語”は教科書以上に問題文に特徴的であると考えられた。

表1 頻出語(相の類)%

抽出語	問題集	試験	教科書
大きい	14.64	10.35	8.79
速い	13.37	8.15	4.88
水平	7.34	7.28	1.78
長い	4.88	5.35	2.25
一定	3.61	2.81	3.27
鉛直	4.72	4.28	1.27
高い	3.33	3.47	2.68
滑らか	2.57	2.87	0.25
垂直	2.24	2.41	1.80
一樣	1.95	2.34	0.65

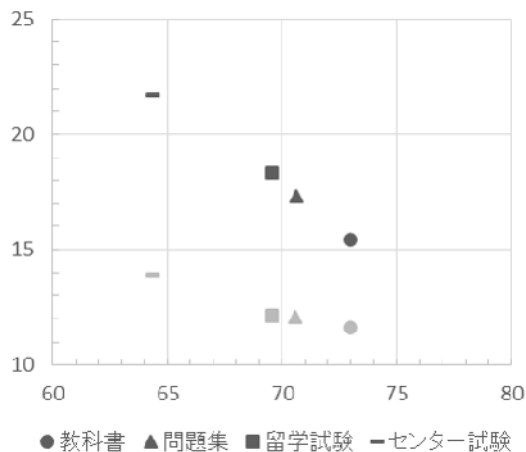


図1 媒体毎の品詞構成比率(異なり語)

表2 問題文の「相の類」特徴語(BCCWJ比)

I【設問に特徴的な一般語】	それぞれ 適当 正しい ところ とき
II【物理に特徴的な一般語】	大きい 高い 上 間 (平板) 時刻 毎秒
III【物理の文脈で専門的な働きをする一般語】	鉛直 水平 垂直 平行 一定 定常 一樣 直後 なめらか あるいは 軽い 長い 速い 静か 十分

表3 問題文の「相の類」特徴語(教科書比)

I【設問に特徴的な一般語】	適当 正しい ところ
II【物理に特徴的な一般語】	(平板) 時刻
III【物理の文脈で専門的な働きをする一般語】	鉛直 水平 一樣 直後 初め なめらか 軽い 長い 厚い 静か

(「平板」はセンター試験の用例においては全て名詞であった)

(3)科目毎の特徴

図2は大学入試センター試験本試験の各科目の問題文をコーパスとして、それぞれの「異なり語」による品詞構成比率を表したものである。自立語を体の類・用の類・相の類・その他の4類に分け、横軸に「体の類」の百分率を取り、「用の類」・「相の類」の百分率を縦軸に取った散布図となっている。図2も図1と同様の特徴として散布図の点が一直線上に並んで見える。これは、体・用・相のどれか一つの比率からテキスト全体の品詞構成が決まることを示しており、“大野-水谷の法則”と一致する。科目毎の文体的な特徴が試験問題に反映されていると考えられ、またこのような広い名詞比率（横軸）の範囲で“大野-水谷の法則”が確認されたことは語彙構造の研究としても意義がある。

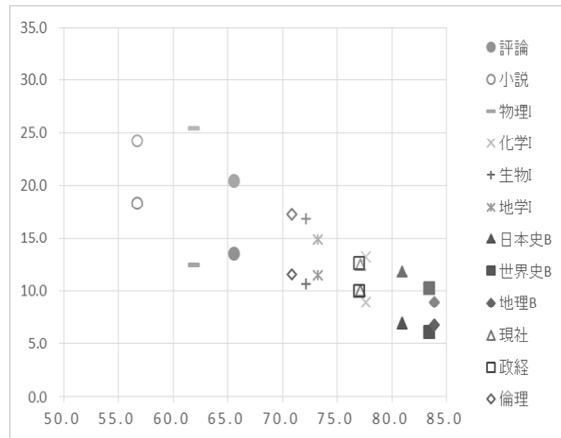


図2 教科毎の品詞構成比率（異なり語）

ここで興味深いのは、物理の名詞比率が小説と評論の中間に位置しており、他の理科の科目（化学・生物・地学）と大きく離れているという点である。先に述べた翻訳語の成立事情も加え、さらに検討する必要がある。

加えて、物理に関しては「延べ語数」・「異なり語数」での品詞構成比率を比較しても、ほとんど変化が見られなかった。これは同一「語」の繰返しや定型表現が少ないことを示唆しており、試験問題文としては特異的である。

(4)教授言語としての日本語

「日本語による専門的なコミュニケーション」を円滑に行うためには「専門語」の習得が欠かせない。本研究では、これまでの専門語研究で十分に捉え切れていなかった、「専門家が無自覚に用いている専門表現」に対して新しいアプローチを提案できた。そしてこれは、

「日本語による専門的なコミュニケーション」を支える日本語の基盤を考えるという意味で、専門語研究に広がりをもたらすであろう。そのために母語話者・非母語話者の別を超えて、日本語で学ぶという過程に付随する“教授言語としての日本語”という枠組みが必要となることを提案した。そして本研究のこれからの広がりには、日本語を母語とせず日本国内で活躍する高度人材を受け入れる箇所まで今後高まると予想される需要に、応えるものになると期待される。

表4 Classification of Vocabulary

	Meaning	Subjects	Specified in	Example
Common Words	General	General	Japanese Dictionaries	台車
Academic Words	Specific	General in Academic	Japanese Dictionaries	観測
Context-Dependent Terms	Specific in Subject	Specific	Practical Situation	なめらか
Technical Terms	Specific in Subject	Specific	Dictionaries of Terminology	速さ

5. 主な発表論文等

〔雑誌論文〕(計 3 件)

[1] 喜古正士・田辺直行 (2018)「ミャンマーの国定物理教科書の分析 力学分野の日本の検定教科書との比較」『日本語教育センター紀要』14, pp.23-32

〔学会発表〕(計 6 件)

[1] KIKO, Masashi (2018) “Japanese as a Medium of Instruction”, 2018 Research Workshop of Institute of Language and Speech Science

[2] 喜古正士 (2018)「物理分野における語彙の構造 テキスト別に見る品詞構成とその特徴」, 第20回専門日本語教育学会研究討論会

[3] 喜古正士 (2017)「日本語の基準としての国語辞書」, 語彙・辞書研究会 第52回研究発表会

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。