

# 医師国家試験の名詞語彙の 対数尤度比に基づく分析と 教材開発の可能性

山元一晃・稲田朋晃・品川なぎさ

## ◆要旨

本稿では、医師国家試験に出現する名詞の分析、および、それを基にした教材開発の可能性について述べた。名詞の分析にあたっては、医師国家試験に特徴的な語を示すため、対数尤度比を用いて、高特徴度語、中特徴度語、低特徴度語に分類した。また、分類した語が上級の日本語学習者にとって既知であるのかを知るため、「日本語教育語彙表」と対照させ、レベルに応じた分類を行った。その結果、低特徴度語は95%以上が一般的な日本語教育で扱われる一方、高特徴度語は30%以下しか一般的な日本語教育で扱われないことが示唆された。次に、教材開発の可能性について述べ、上記の情報のほかに語の分類など様々な指標が必要になることを指摘した。

## ◆キーワード

医師国家試験、医学用語、コーパス、語彙教材、専門日本語

## ◆ABSTRACT

This study analyzes the nouns used in the National Examination for Medical Practitioners and discusses the development of related language teaching materials for international medical students. To statistically analyze the nouns, we calculated the log-likelihood ratio of each word and divided them into three groups: high-characteristic, middle-characteristic, and low-characteristic words. We then compared the list of words with *Nihongo Kyoiku Goiho* [the Japanese Learners' Dictionary] (Sunakawa et al, 2012) to identify the words that an advanced learner of Japanese may or may not be able to understand. The results showed that among the high-characteristic words, less than 30% may not be taught in Japanese language classes. Based on the analysis, we conclude that each word requires several labels according to its various characteristics, in order to develop useful language teaching materials.

## ◆KEY WORDS

National Examination for Medical Practitioners, medical terms, corpus, teaching materials, technical Japanese

## Log-likelihood Analysis of Nouns in the National Examination for Medical Practitioners and the Development of Teaching Materials for International Medical Students

KAZUAKI YAMAMOTO, TOMOAKI INADA,  
NAGISA SHINAGAWA

## 1 研究の背景と目的

日本において医師を目指す留学生は一定数いるが（松元 2000, 守山 2004 など）、今後はそのような留学生がますます増えることが予想されている。たとえば、2017年に医学部医学科を新設した首都圏のある大学では、初年度から毎年20名ずつの留学生が入学し、6年生までが揃う2022年度には、約120名の学生が在籍することになる。彼らの大半は、日本語を学びながら日本の医師免許の取得を目指す。

一方で、医師を目指す留学生のための日本語教育に資する研究の蓄積はほとんどなく、基礎研究および教材の開発が必要である。なかでも、日本国内で医師として従事するために必須の資格であり、かつ、難度の高い語彙が使用される医師国家試験への対応は急務である。そこで、本稿では、基礎研究として、以下の2点を明らかにする。

- ①医師国家試験における特徴的な語彙とは何か
- ②医師国家試験に特徴的な語は、一般的な日本語教育で扱われるのか

なお、研究対象として医師国家試験を用いるのは、日本国内で医師として従事することを目指す全ての学生が受験するためである。

## 2 先行研究

先述のとおり、医師を目指す留学生のための研究はほとんどない。一方で、看護師、介護福祉士については、2008年から経済連携協定（EPA）に基づいた候補者の受け入れが始まったこともあり、看護、介護分野での調査、報告は多い（岩田 2014, 大場 2017, 奥田 2011 など）。以下に看護師国家試験の分析の先行研究として岩田（2014）を、介護福祉士国家試験の分析の先行研究として大場（2017）を挙げる。

岩田（2014）は、看護師国家試験（過去7回分）の文法、語彙を旧日本語能力

試験の出題基準と比較し分析している。文法については、1級文法は看護師国家試験にはほとんど出現しないことを明らかにしている。一方、語彙については、抽出した名詞語彙を旧日本語能力試験のレベルと対応させた結果、65%以上が1級・級外の語彙であり、2級レベルでは全く対応できないと主張している。

大場（2017）は、介護福祉士国家試験（過去6回分）の文法、語彙を分析し、旧日本語能力試験出題基準を用いたレベル判定を行っている。文法については1級文法は1例のみであり、2級文法も限られた項目のみが頻出していることを指摘している。語彙については、名詞が延べ・異なり語数ともに動詞と比較して圧倒的に多く、名詞語彙は級外語彙を含む多数の専門語彙が出現することを明らかにしている。

上述した2つの先行研究は、国家試験に出題される専門語彙は一般的な日本語教育では扱われない可能性があり、その学習方法などについて検討する必要性を示唆している。

## 3 分析の方法

先行研究を参考にしながら、まずは医師国家試験問題を文法面、語彙面から多角的に分析し、その上で医学部留学生のための日本語教育の在り方を検証していく必要がある。本稿ではそのうち語彙面、特に名詞を中心に分析する。

まず、「医師国家試験に特徴的な語は何か」という点を明らかにするために、医師国家試験にどの程度特徴的であるかを示す指標を定め、その指標に基づいて医師国家試験に特徴的な語を客観的に示し、特徴度ごとに分類する。

次に、「医師国家試験で用いられる語彙は、一般的な日本語教育で扱われるのか」を明らかにするために、分類した語を「日本語教育語彙表」（Sunakawa et al. 2012）<sup>[註1]</sup>と対照する。そして、日本語学習者が学ぶ機会があるか、また、あるとすれば、どのレベルの学習者であれば既知と考えられるのかを明らかにする。

また、5節において上記の医師国家試験の分析を基にした教材開発の可能性

について考察する。

### 3.1 分析に用いるデータの概要

本研究では、医師国家試験のうち、本稿執筆時最新の6回分、106回（2012年）から111回（2017年）をデータとして用いる。医師国家試験は、各年問題A～問題Iがあり、それぞれが、問題冊子と別冊に分かれている。別冊は、写真や図を中心としたものであるため分析対象とはしなかった。また、問題冊子についても、問題文および選択肢に含まれるもののみ対象とした。上記をテキストデータ化し、形態素解析を行った後、空白・空行を機械的に除去し、名詞として分類されていた記号（「」など）について目視により確認、修正を行った。

なお、形態素解析エンジンにはMeCab 0.996を用いた。また、形態素解析辞書については、国立国語研究所の「現代日本語書き言葉均衡コーパス」（以下BCCWJとする）を対照コーパスとして用いるためUniDic 2.1.2を用いた。

異なり語数は9,237で、そのうち名詞は7,629語（82.6%）であった。そのうち漢字・ひらがな・カタカナのいずれかを含む名詞6,571語（71.1%）を分析対象とした。

### 3.2 特徴語抽出の方法

英語教育の分野では、特定のコーパスに特徴的な語を抽出するために、様々な指標が用いられている。そのうち、対数尤度比（Log-likelihood ratio, Dunning 1993）は、比較的小さなサンプルでも有効であることが指摘されている（Chujo & Utiyama 2006など）。また、日本語教育においても、寺嶋（2009）が日本語教育語彙を選定するための統計的指標として、対数尤度比を用いた尤度比検定、カイ2乗検定、イエーツの補正公式を比較している。その結果、3つの指標のうち、尤度比検定はコーパスサイズに最も影響を受けなかったとしている。その他、BCCWJを基にした「教科特徴語リスト」<sup>[註2]</sup>は、対数尤度比を用いて、各教科に特徴的な語を抽出している（近藤 2011）。医師国家試験は、毎回内容の精査が行われ新たな知見が追加されたり、古い情報が削除されたりするため、むやみにコーパスサイズを大きくすることは適切ではない。よって、コーパスサイズに最も影響を受けないとされる対数尤度比を本分析の特徴度の指標とする。

対数尤度比は、その値が高ければ高いほど、どちらかのコーパスに特徴的な語であるといえる。対数尤度比（ $G^2$ ）は、以下の式により求められる（近藤 2011）。

$$G^2 = 2(\log(a)+\log(b)+\log(c)+\log(d)-(a+b)\log(a+b) - (a+c)\log(a+c)-(b+d)\log(b+d)-(c+d)\log(c+d)+(a+b+c+d)\log(a+b+c+d))$$

a : 当該テキストでの当該語の頻度  
b : 参照テキストでの当該語の頻度  
c : 当該テキストの当該語以外の頻度  
d : 参照テキストの当該語以外の頻度

ただし、 $\log(x)$ は自然対数を表す。

対数尤度比は、当該テキストに比して参照テキストに特徴的である語であっても高い値となるため、 $ad-bc < 0$ の場合、-1を乗じる補正を行い、特徴度とする。正の値の特徴度の有意水準と臨界値は表1の通りである。

表1 正の値の特徴度の有意水準と臨界値

有意水準	p<.05	p<.01	p<.001	p<.0001
臨界値	3.84	6.63	10.83	15.13

対照用のコーパスには、BCCWJを用いた。1億430万語とコーパスサイズが大きく、かつ、分野のバランスが取られているためである。

医師国家試験には医学用語が多数含まれており、BCCWJの一部のみを対象にすると、頻度が0になる語が多数となる可能性がある。そのため、特徴語を算出する際に、BCCWJの全コーパスを対象とした。

### 3.3 特徴度算出の対象とする語

形態素解析によって抽出された名詞のうち、BCCWJの各語彙素の頻度データである「短単位語彙素表データ」にも含まれている6,370語について特徴度を算出した。上記に含まれていない201語については、以下の可能性がある。

①形態素解析において、語の途中で区切れているなどの語（例：ンション）

②BCCWJに実際に含まれていない語（例：鶯卵、蜂窩）

上記のうち①については、目視により形態素解析結果を修正していくことで、ある程度解決できる。また②については、頻度が20の語も含まれていることを踏まえ、医学の専門家の意見を参考にしながら、必要な語を抽出する必要がある。これについては4.4で詳しく述べる。

## 4 特徴度の分析結果

対数尤度比に基づく特徴度によって、以下のように語彙を分類する。

- ・高特徴度語：有意に特徴度が高い語 ( $p < .05$ )
- ・中特徴度語：高特徴度語にも低特徴度語にもあてはまらない語
- ・低特徴度語：有意に特徴度が低い語 ( $p < .05$ )

分析対象とする6,571語中、「高特徴度語」に該当する語は2,894語（44.0%）、「中特徴度語」に該当する語は2,228語（33.9%）、「低特徴度語」に該当する語は1,248語（19.0%）であった。またBCCWJに含まれない語は201語（3.1%）だった。

### 4.1 各分類に該当する語

表2は、高特徴度語のうち、特徴度が高いもの10語と低いもの10語を示している。医学に関連する語が多いことが分かるが、「呼吸」など日本語学習者でも知っていると考えられる語が含まれている一方で、「主訴」などBCCWJではほとんどみられない語も含まれている。これらについては、「日本語教育語彙」と比較して、取捨選択ができると考えられる。また、「球」など、本来は「白血球」「赤血球」などの一部となっている語も含まれている。

医師国家試験での頻度が低い語が高特徴度語として分類されているものもある。これは、BCCWJでの頻度が極めて低いためであり、これらの語については検討の余地がある。

表2 高特徴度語のうち特徴度が高い10語および低い10語

特徴度が高い語				特徴度が低い語			
順位	語彙素	頻度	特徴度	順位	語彙素	頻度	特徴度
1	所見	1,562	14,316.7	2,885	麻薬	9	3.92
2	来院	1,146	12,138.7	2,886	受け答え	2	3.92
3	別冊	1,049	10,515.8	2,886	同大	2	3.88
4	球	1,099	9,234.2	2,888	たらの芽	1	3.88
5	主訴	830	9,035.4	2,888	ビレン	1	3.88
6	呼吸	1,147	7,450.6	2,888	顎骨	1	3.88
7	血圧	1,006	7,359.5	2,888	去痰	1	3.88
8	整	663	7,351.6	2,888	経陸	1	3.88
9	異常	1,165	7,170.4	2,888	吹き込み	1	3.88
10	血液	1,062	6,966.1	2,894	鏡面	2	3.85

表3は、低特徴度語のうち、特徴度が高い10語および低い10語を示している。これらの語については、一般的な語がかなりの程度含まれており、除外しても問題はないと考えられる。ただし、低特徴度語であるにもかかわらず日本語教育語彙表に含まれていない語もあり、これらについては4.3で考察する。

表3 低特徴度語のうち特徴度が高い10語および低い10語

特徴度が高い語				特徴度が低い語			
順位	語彙素	頻度	特徴度	順位	語彙素	頻度	特徴度
5,123	利害	1	-3.84	6,362	六	1	-1,059.40
5,124	床	11	-3.84	6,363	者	27	-1,096.91
5,125	復旧	1	-3.85	6,364	四	12	-1,319.79
5,125	引越	1	-3.85	6,365	物	76	-1,339.90
5,127	先進	4	-3.87	6,366	事	839	-1,511.18
5,128	総合	27	-3.87	6,367	五	1	-1,558.24
5,129	工場	16	-3.87	6,368	三	25	-1,880.52
5,130	桁	1	-3.88	6,369	十	8	-1,982.77
5,131	執筆	1	-3.89	6,370	二	60	-2,234.61
5,131	畳	1	-3.89	6,371	一	78	-2,806.25

医師国家試験に有意に特徴的でないとされる語である中特徴度語についても、一部に、医師であれば必要な語が含まれている。例えば、中特徴度語に分類される「卵黄」は、(1)のように用いられている。

- (1)
- 3 初期離乳食に関する母親への指導で適切なのはどれか。
- a 開始時期は生後7か月とする。
  - b 卵は卵黄のみを用いる。
  - c 穀類は粥ではなく軟飯を与える。
  - d 硬さは歯茎でつぶせる程度にする。
  - e 人工乳栄養児の場合は牛乳に変更する。

(第106回(2012年)問題G ※下線引用者)

この問題において、「卵黄」という語の意味が分からなければ、問題に答えることは難しい。有意に特徴的である語のみを提示することで良いのかは検討の余地がある。これについては、5.3でも述べる。

## 4.2 「日本語教育語彙表」との対照

特徴度が有意に高い語が、日本語教育で一般に扱われる語彙に含まれるのかどうかを検討するため、「日本語教育語彙表」を参照した。それぞれの語がどのレベルに該当すると考えられるのかを表4に示す。なお、「日本語教育語彙表」に準じ、一般的な表記に直して対照した。また、「赤血」「血球」のように元の語(赤血球)が容易に推測でき、かつ、「日本語教育語彙表」に該当の語がある場合は、該当ありとした。

全異なり語6,571語のうち、「日本語教育語彙表」の初級前半～上級後半に含まれる語は、3,324語(50.6%)であり、該当なしの語は、3,247語(49.4%)であった。

また高特徴度語の2,894語のみに絞ると「日本語教育語彙表」の初級前半～上級後半に含まれる語は787語(27.2%)であり、該当なしの語は、2,107語(72.8%)であった。特徴度が有意に高い語は、「日本語教育語彙表」に含まれない

表4 特徴度ごとの日本語教育語彙表との対応

		初級前半	初級後半	中級前半	中級後半	上級前半	上級後半	該当なし	合計
高特徴度語	語数(語)	17	42	99	302	257	70	2,107	2,894
	割合(%)	0.6	1.5	3.4	10.4	8.9	2.4	72.8	100.0
中特徴度語	語数	24	71	191	583	409	54	896	2,228
	割合(%)	1.1	3.2	8.6	26.2	18.4	2.4	40.2	100.0
低特徴度語	語数	106	148	340	456	147	8	43	1,248
	割合(%)	8.5	11.9	27.2	36.5	11.8	0.6	3.4	100.0

語が大半を占めることがわかる。

一方、低特徴度語については1,248語のうち「日本語教育語彙表」に含まれる語は1,205語(96.6%)であり、該当なしの語は、43語(3.4%)であった。低特徴度語は、そのほとんどが、「日本語教育語彙表」に含まれる。

## 4.3 低特徴度語のうち日本語教育で扱われないとされる語

本節では、低特徴度語のうち、「日本語教育語彙表」に含まれていない語について、具体的に見ていきたい。表4のとおり該当する語は43語であった。その多くが頻度2以下であるが、頻度が3以上のものもいくつかあり、これらを(2)に示す。そのうち「中国」は身近な固有名詞であり、「バス」は「日本語教育語彙表」に含まれてはいないが初級の段階で学ぶと考えられる。なお、( )内の数字は頻度を表す。

(2) ホーム(18) 中国(7) バス(6) シート(5) 助成(3)

いずれも、医学用語とは考えにくい語であるが、頻度の高い「ホーム」についてその用例をみていきたい。「ホーム」については、18例のうち、「特別養護老人ホーム」等「老人ホーム」に含まれている例が13例、「グループホーム」に含まれている例が3例、「ホームヘルパー」に含まれる例が1例であり、英語の“home”の借用語がほとんどであることが分かる。これらの語は、複合語

としては医療や福祉に独特の言葉であり考慮する必要がある。

#### 4.4 BCCWJに含まれていない語

次に、医師国家試験には含まれているが、BCCWJには含まれていない語について考察したい。前述したように、異なり語数で201あり、延べ語数で451あった。本稿で対象とした名詞の全延べ語数に占める割合は、0.33%である。そのうち、頻度が2以上のものは、85例ある。紙幅の都合上頻度5以上の語を(3)に示す。なお、( )内の数字は頻度を表す。

- (3) 開大 (29) カニューラ (20) リウマトイド (16) 鼻閉 (11)  
フロセミド (10) アムホテリシン (9) ハプトグロビン (8)  
妊時 (8) コルポスコピィ (7) 尖弁 (7) クリプトコックス (6)  
サイアザイド (6) プロラクチノーマ (6) ポリソムノグラフィ (6)  
促迫 (6) アドヒアランス (5) ビグアナイド (5) ヘプシジン (5)  
両下 (5)

頻度が5以上の語のみを見てもカタカナ語が多いことがわかる。これらの語は、BCCWJに含まれておらず、一方で、医師国家試験には一定数含まれており、覚えておかなければならない言葉であると思われる。また、頻度4以下の語もカタカナ語が多く、漢字からの類推ができないため、覚えなければならないと考えられる。

ここでは、頻度が最も高い「開大」が用いられている文脈(4)をみたい。

- (4)  
31歳の2回経妊1回経産婦。妊娠39週2日。陣痛発来のため入院した。分娩は順調に進行し、入院3時間後に子宮口8cm開大、未破水である。規則的な陣痛を認めており、胎児心拍数パターンに異常を認めない。

(第107回(2013年)問題G ※下線筆者)

このように「開いている」という意味で用いられている。

また、頻度が2番目に高い「カニューラ」については、いずれも「鼻カニューラ」として用いられている。

これらのような語は医師国家試験においては注釈なしで用いられている。したがって、頻繁に用いられているにもかかわらずBCCWJの全コーパスに見られない語については、リストアップしておく必要がある。

## 5 教材開発の可能性

高特徴度語の多くが、一般的な日本語教育では扱われないことを考えると、医師国家試験を目指す留学生のために早急に教材開発に向けた語彙リストを作成する必要がある。さらに、特徴度という指標を用いて、医師を目指す留学生が学ぶ必要のある語を抽出し、その中から、さらに日本語教育で学ぶかどうかを指標に分類し、教材作成へと活かしていく必要がある。

本節では、4節で述べた分析結果を基にした語彙リスト提供の可能性、およびそれを基にした、教材開発のためのいくつかの方向性を述べたい。

### 5.1 医師国家試験語彙リストの提供

本項では、教材開発の基礎となりうる語彙リスト作成の計画について述べる。リスト作成にあたっては具体的には、以下の情報を付与する。

- ①特徴度による分類(高特徴語、中特徴語、低特徴語)
- ②「日本語教育語彙表」による分類
- ③頻度
- ④各語を含む複合語の情報およびその頻度
- ⑤英訳
- ⑥医師国家試験を引用した例文

上記のうち①～③については、既に本研究で結果が得られている。④については、本研究で行ったような分析方法では一括して複合語を抽出できないため、すべての語を逐一調べていく必要がある。また、⑤～⑥については、専門

家の意見を聞きながら作成する必要がある。これらについては、その方法を今後精査し、語彙リストの作成を進めたい。

## 5.2 レベルに合わせた高特徴度語の学習の可能性

医師国家試験合格を目指す中級レベルの学生への日本語教育に活かせる語彙として、高特徴度語でありながら、初級および中級レベルとして分類されている語があげられる。(5)は、特徴度が上位20位以内に含まれていながら、初級前半～中級後半に該当する語である。

(5) 分、球、呼吸、血圧、異常、血液、胸部、検査、基準、体温

上記のうち「分」は、「脈拍84／分」のように使われており、特に医学的に特徴的な使い方であるとは言い難い。しかし、たとえば、「呼吸」や「異常」のように語としては中級学習者であれば理解していると考えられる語であっても、(6)のように、動詞一般に分類される語の中で最も高頻度かつ特徴度が高い動詞（「認める」）とともに用いられている例がある<sup>[註3]</sup>。

(6) 呼吸音に異常を認めない。 (第107回(2013年)問題A)

上記の文では、いずれの語も「日本語教育語彙表」では中級後半までに学ぶとされている。このような例が散見されることから、中級までの段階で学習することのできる特徴的な語が含まれる文を例文として提示することで、早い段階であっても医師国家試験でよく用いられる表現に慣れることができると考えられる。

また、(7)に「胸部」を含む複合語の例を示す。

(7) 胸部エックス線、胸部不快感、胸部中部進行食道癌、胸部単純CT、胸部圧迫感、胸部放射線放射、胸部異常陰影、胸部痛、胸部絞扼感、胸部造影CT

「胸」という漢字を学んだ際に、その他の語の構成要素のレベルを考えた上で、「胸部」を含む複合語も同時に学べるはずである。高特徴度語を含む語をリストアップしておけば、効率的に関連語も学べると考えられる。また、上級においても、同様の方法で、それぞれのレベルにあった語を効率よく習得できると考えられる。

## 5.3 高特徴度語かつ日本語教育語彙表に含まれていない語の学習の可能性

5.2では、医師国家試験に特徴的であると考えられる名詞を日本語教育での各段階で導入する方法を提案したが、本項では、高特徴度語の70%以上にのぼる「日本語教育語彙表」に含まれていない語の教材についての提案をする。

高特徴度かつ「日本語教育語彙表」に含まれない語の例として、特徴度が上位20位以内の語を(8)に示す。

(8) 所見、来院、主訴、整、脈拍、腹部、投与、小板、触知、清明、クレアチニン、好中、リンパ、尿素、アルブミン、心音、結膜、既往、造影、浮腫

「所見」、「主訴」などは、症例問題で定型的に用いられる語であり、「整」は、1文字で不整脈でないことを表すものである。このような語は、その語の意味が一度分かっただけで問題になることはないと思われる。「清明」も「意識が清明である」のように定型的に用いられており、典型的な例文を示すことで、理解ができるように思う。

また、「来院」「投与」「触知」は「する」を伴って動詞として用いられ、実際に本文では動詞として用いられている例がほとんどである。このような語は、医師のすること、患者のすることできく分類できる。

一方、「腹部」「結膜」は身体の一部を表すものである。「クレアチニン」「尿素」「アルブミン」は物質名であるし、「小板」「好中」などは、それぞれ「血小板」「好中球」などのように語の要素として用いられるものである。「浮腫」のように症候を表すものもある。上記のようにいくつかの分類を行い、学生に提示することができるだろう。

また、それぞれの語は、日本語教育で学ばないとしても、たとえば、「血小板」のような語であれば、それぞれの漢字は初期の段階で学ぶと考えられるため、漢字学習に取り入れることもできる。

5.1および5.2で述べたように、特徴度および日本語教育におけるレベルで分類することで、日本語教育の各段階で取り入れることができ、さらに、それぞれの語を分類することで教材化が容易になってくると考えられる。

## 6 まとめと今後の課題

### 6.1 本稿のまとめ

本稿での分析は、以下の3点にまとめられる。

- ①医師国家試験に特徴的であると考えられる語は2,894語あった。
- ②医師国家試験に特徴的であると考えられる語の70%以上は、一般的な日本語教育では扱われないことが示唆された。しかし、見方を変えれば、特徴度が上位の語であっても30%程度は日本語教育でも扱われるとも考えられ、これらは日本語の習得状況に合わせて指導することができる。
- ③特徴度で分類し、日本語教育でのレベルを付与した語彙リストを提供することで、医学用語の学習のための教材開発に活用できる可能性を示した。

### 6.2 今後の課題

今後の課題として、以下の3点があげられる。

- ①他品詞の分析も必要である。特に、名詞には「する」を伴ってサ変動詞になるものがあるため、動詞の分析が必要である。
- ②本稿の分析では短単位による形態素解析しか行っていないため、複合名詞の分析はできていない。医師国家試験には複合語が多く用いられているため、複合語の分析が必要である。
- ③教材作成やリストの提供にあたっては、「英訳」「例文」など、より多くの

情報を付与する必要がある。

- ④語彙リストを基にした教材開発を行い、医師国家試験を目指す留学生のための教材開発を進めていく必要がある。

①については、今後、同様の方法を用いて分析を行っていく。②については、高特徴度語を含む複合語をピックアップしリスト化していく。③については、医師等医療従事者の協力を仰ぎながら、必要だと考えられる情報を付与していきたい。④については、現在、医師国家試験に用いられている上級レベルの漢字を抽出した教材開発を試みている。その際に、本語彙リストを、漢字ごとに提示する語彙を客観的に抽出するために使用する予定である。この教材については、稿を改めて紹介したい。 <山元・稲田・品川 いずれも国際医療福祉大学>

### 付記

本研究では、日本語教育語彙表 ver1.0 (<http://jhlee.sakura.ne.jp/JEV.html>) を利用した。

### 注

- [注1] …… 「日本語教育語彙表」は李在鎬氏のウェブサイト (<http://jhlee.sakura.ne.jp/JEV.html>) でダウンロードできる。
- [注2] …… 国立国語研究所のウェブサイト ([http://pj.ninjal.ac.jp/corpus\\_center/bccwj/freq-list.html](http://pj.ninjal.ac.jp/corpus_center/bccwj/freq-list.html)) よりダウンロードできる。
- [注3] …… 動詞についての分析は紙幅の都合上紹介できないため、稿を改めたい。

### 参考文献

- 岩田一成 (2014) 「看護師国家試験対策と「やさしい日本語」」『日本語教育』158, pp.36-48. 日本語教育学会
- 大場美和子 (2017) 「介護福祉士国家試験の筆記試験における文法・語彙項目の分析—日本語能力試験の観点から」『小出記念日本語教育研究会論文集』25, pp.5-20. 小出記念日本語教育研究会
- 奥田尚甲 (2011) 「看護師国家試験の語彙の様相—日本語能力試験出題基準語彙表との比較から」『国際協力研究誌』17(2), pp.129-143. 広島大学大学院国際協力研究科
- 近藤明日子 (2011) 「中学校・高校教科書の教科特徴語リストの作成」『特定領域研究「日本語コーパス」言語政策班報告書 (JC-P-10-01)』pp.145-152. 国立国語研究所

- 寺嶋弘道 (2009) 「日本語教育語彙を選定するための統計的指標」『ポリグロシア』17, pp.71-83. 立命館アジア太平洋研究センター
- 松元宏行 (2000) 「医学系留学生のための専門日本語教育」『専門日本語教育研究』2, pp.46-53. 専門日本語教育学会
- 守山恵子 (2004) 「医学系留学生のための日本語クラスの試み」『長崎大学留学生センター紀要』12, pp.1-13. 長崎大学留学生センター
- Chujo, K. & Utiyama, M. (2006) Selecting level-specific specialized vocabulary using statistical measures. *System*, 34, pp.225-269.
- Dunning, T. (1993) Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1), pp.61-74.
- Sunakawa, Y., Lee, J., & Takahara, M. (2012) The construction of a database to support the compilation of Japanese learners' dictionaries. *Acta Linguistica Asiatica*, 2(2), pp.97-115.