コロケーション検索システム「かりん」の開発と試行
―中上級日本語学習者の試行結果を比較して―

中　溝　朋　子
坂　井　美恵子
金　森　由　美

要旨

本稿では，モバイル端末向けに開発したコロケーション検索システム「かりん」の特徴について簡単に紹介した後，2019 年度に実施した中上級レベルの日本語学習者の試行結果を述べ，本システムの今後の課題について考察する。

キーワード

コロケーション，コーパス，中上級日本語学習者，機能動詞

1.　はじめに

従来，日本語教育におけるコロケーション，や連語の重要性は指摘されてきたが，近年，日本語におけるコーパスの開発が進むにつれ，その成果を日本語学習に活用する教材が活発に作成されるようにになった。一方でコーパスを活用したコロケーションの検索システムも「NINJAL-LWP for BCCWJ」[1)]，「NINJAL-LWP 筑波 web コーパス」[2)]，日本語作文支援システム「なつめ」[3)] などが開発されている。これらのシステムでは，様々なコーパスを用いて抽出されたコロケーションが統計指数別に示され，例文も併せて閲覧可能となっている。しかし日本語学習者にとっては①コロケーションが一度に多数表示されるため，コロケーションの選択が難しいこともあること，また②表示される例文がコーパスの原文そのままであるため，文脈の不足，漢字の読み・語彙・文法の難しさなどの点で理解が容易ではない場合があるなど，学習者が実際に使用するには難しい面もあった。さらにこれらはタブレットやスマホなどのモバイル端末への対応版はまだ行われていない。

これらを踏まえ，筆者らは，中上級日本語学習者を対象とした「日本語学習者のためのコロケーション検索システム「かりん」（モバイル版）を開発した（以下，「かりん」）。本稿では，「かりん」について，簡単に概要，および特徴を操作方法を紹介しながら述べた後，2019年度に実施した試行の結果を述べ，今後の改善点・課題について検討する。

2.「かりん」の概要

2.1 データ

「かりん」では，「現代日本語書き言葉均衡コーパス」（国研2011，以下，BCCWJ）より，主に旧日本語能力試験（以下，旧JLPT）１級および２級の名詞と，それらと共起する語（修飾語と動詞）を抽出したデータを用いている。BCCWJ は「現代日本語の書き言葉の全体像を把握するために構築したコーパス」であり，「書籍全般，雑誌全般，新聞，白書，ブログ，ネット掲示板，教科書，法律などのジャンルにまたがって１億 430 万語のデータ

を格納」している（国研2011）。「かりん」では，上述の名詞，およびこれらと共起する修飾語と動詞のコロケーションを正規表現によって検索・抽出し，共起頻度，および共起強度を示すDice係数[4]を計算し，その結果を表示している。表示する範囲は，共起頻度５以上とし，画面にはDice係数順に表示している。

3.「かりん」の特徴

「かりん」はタブレットやスマホなどのモバイル端末に対応したシステムである。対象は，従来のコロケーション検索システムより広いレベルの学習者，具体的には日本語能力試験（以下「JLPT」）のN2程度の学習者から使用が可能になることを目指している。そのために以下のような点に留意をした。

(1) デザインや機能をシンプルにし，直観的に操作できるようにする
(2) 検索機能を，学習者が最も必要であろうとするコロケーション「名詞（＋助詞）＋動詞」，「修飾語＋名詞」の検索に絞る。すなわち中心語として名詞を入力し，その共起語である「動詞」もしくは「修飾語」を検索可能とする。このように機能を限定している半面，学習者に必要な機能を提供する。
(3) 学習者のための機能としては，①操作ボタンの２言語対応（日英），②ふりがなの表示，③JLPTや学習指標値（徳弘2006）など語のレベルや学習の重要度の表示，④村木（2000）の「機能動詞」を基に分類した色による意味分類の表示，⑤作例による例文・翻訳（中・韓・英）の表示，⑥コロケーション別にBCCWJのサブコーパス別の使用頻度の表示などである。

以下，これらの機能について簡単に説明する。

４．「かりん」の操作方法

以下，「かりん」の操作方法を紹介しながら，3.で述べた学習者機能について簡単に紹介する。

図１は「かりん」のトップ画面である。ここで操作ボタンの言語を日英から選び，「確認事項」を読んだ後，検索が可能となる。

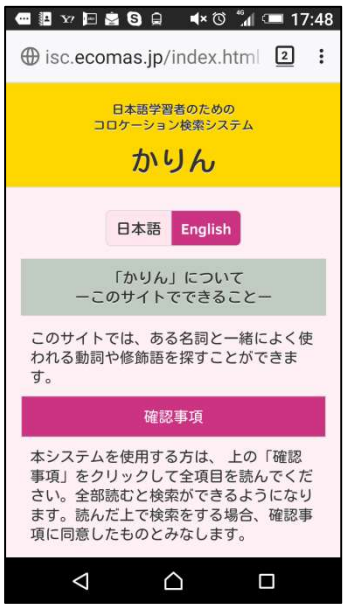

図１「かりん」トップ画面

検索を開始する前に，図２のように検索する「名詞のレベル（JLPTのレベルによって選択）」，および検索結果の訳の言語を中韓英の中から選択する図２のようなカスタマイズ機能もある。

検索は，図３の画面からまず「名詞＋動詞」もしくは「修飾語＋名詞」を選択した後，名詞を入力して検索する。

図2 表示のカスタマイズ画面

図3 検索開始画面

図4 名詞入力画面（直接入力）

図5 名詞入力画面
（名詞リストから選択）

名詞の入力の仕方は，図4のような直接入力と図5のような「名詞をリストから選ぶ」方法のから選択する。

直接入力の場合は，文字を入力するにつれ，収録語の中から候補を　する　機能も付与し，入力の労力の軽減を図っている。

「名詞をリストから選ぶ」を選択すると，図5のような 50 音表が現れるので，先頭文字を選ぶと収録する語のリストが現れる。

図6は，「ひらがな」の「か」を選択した場合に現れるリストである。名詞の右側には，当該名詞の JLPT のレベルと学習指標値が示されている。この中から検索したい語（例えば「開発」）をタップすると図7のような検索結果が現れる。

図6 先頭文字をタップした後の画面
（例：先頭文字「か」）

図7「開発」の動詞コロケーション検索結果

検索結果は，ふりがな，動詞の JLPT レベル - 学習指標値（各動詞の下），BCCWJ における頻度とそれを基に計算した Dice 係数とともにコロケーションが示される。さらにコロケーションをタップすると図 8 のような例文と翻訳画面が現れる。

図8 例文と翻訳画面

図9 サブコーパス別相対頻度画面

この画面をさらにスクロールすると図 9 のような BCCWJ のサブコーパス別の相対頻度がグラフで表示される。これにより学習者は，コロケーションの出典がわかり，実際に使用する際に参考にすることができる。

図 10 機能動詞意味の色による分類説明画面

また図 7 のコロケーションと頻度の帯グラフの間の 2 つの色は，村木(1991)の「機能動詞」を基に筆者らが行った動詞の分類を示す。左側の色は「アスペクト」「ヴォイス」「ムード」「複合助辞」を表し，右側がさらにその細分類された意味を表す。例えば，図 10 のように「アスペクト」は，「開始」「変化・出現」などに分かれており，同じ意味に属するものは，2 色ともに同じ色で示されている。そのため，検索結果画面で多くのコロケーションが示されても，同じ色が示されていれば類義語であることがわかり，学習者のコロケーションの選択に役立つと考えた。

以下，このような特徴を持つ「かりん」を日本語学習者に実際に施行した結果について述べる。

5. 2019 年度「かりん」試行結果について

「かりん」の試行では，日本語中級クラスと上級クラスで実際に学習者に使用してもらい，その結果や感想・改善点を記入してもらった。

以下，この試行の概要と結果について述べる。

5.1 試行の概要と方法

今回の試行の概要を以下，表 1 に示す。

表1「かりん」試行実施概要

	A 大学	B 大学
時期	2019年12月	2020年 1 月
レベル	中級 (N3合格程度)	上級 (N1合格程度)
人数	24名	20名
学習者母語	約半数が中国語。韓国語，その他アジア言語，英語以外欧州言語	主に中国語。韓国語，英語以外欧州言語

各レベルの中国語圏の割合は中級は46%，上級は 80% で，その他の言語は各レベルそれぞれ 1 ～ 5 名程度である。「かりん」はN2程度のレベルの学習者を対象にしており，N3に合格し，N2を目指している本中級クラスの学生は，「かりん」を使い始めることが想定されるレベルである。

今回の試行の目的は，①学習者のこれまでの作文を書く際の語の検索方法（web 辞書など）に比べ，「かりん」を用いることで，より多くの動詞を選べるようになるか，さらに②適切な動詞をより多く選べるようになるか，③中級と上級で効果や使用の方法や動詞の選択の傾向，感想に違いはあるかの 3 点である。

試行の実施方法は，表 2 の 5 つの例文について，文の意味を考えて正しいと思う動詞を書くという内容である。各例文には 3 語記入れきる回答欄を設けたが，回答数は限定しないこととした。1 回めはスマホなど普段作文を書くときと同様の方法で記入してもらい，用紙を回収後，同じ例文について「かりん」を使用して記入してもらった。回収後，「かりん」を使用した感想や改善点に関するアンケートを記入してもらった。

表2 試行回答用　例文

①日本の文化は，古くから中国や韓国の影響（　　）＿＿＿＿＿＿＿＿＿＿。
②ほとんどの自動車会社では，自動運転技術の開発（　　）＿＿＿＿＿＿＿＿＿＿が，まだ，100%自動運転の車はない。
③私は今年の10月，胸に期待と不安（　　）＿＿＿て，日本に留学してきた。
④先月の台風は，西日本に大きな被害（　　）＿＿＿＿＿＿＿＿＿＿。
⑤私の説明が終わったら，作業（　　）＿＿＿＿＿＿＿＿＿＿ください。

以下，試行の回答結果，および試行後に実施したアンケート結果について述べる。

5.2　試行，およびアンケート結果と考察

5.2.1　試行結果

以下，表3に「かりん」の使用の有無による回答数，正答数，異なり語数の違いをレベル別に示したものを示す。

表4　レベル別　回答数・正答数

	中級		上級	
「かりん」使用	前	後	前	後
回答数平均	1.6	2.3	2.0	2.9
正答数平均	2.7	6.5	6.1	9.2
異なり語数計	70	75	70	56
正答異なり語数計	21	29	27	30

表4の「回答数」「正答数」は，1例文当たりの1人当たりの平均数を示している[5)]。「異なり語数」は，記入された動詞の数が全体で何種類出されたか，正誤や記入数に関わらず合計した数で，「正答異なり語数」は，そのうちの正答数の合計数である。

まず，回答数，および正答数，異なり語数については「かりん」を使用することで両レベルともに増えていることがわかり，多様な表現を選択し，またその正答率も向上していることがわかる。

一方，上級では正答の異なり語数は増えているが，誤答も含めた異なり語数は減っている。これは特に明らかな誤答（後述する誤答②）が「かりん」を使用した後に減少していることによると思われる。このように「かりん」を用いることで従来学習者が行っていた方法では，見つけられなかったコロケーションを見つけられたり，自己訂正したりする機会となったのではないかと推測される。

以下，さらに詳しく，表5から表9に例文ごとの正誤別に異なり語や回答数を示し，正答の特徴や誤答の原因について分析する。なお，表中，左列「前／後」は「かりん」の使用前と使用時を示し，動詞後の（　　）内の数字は回答数を示す。学習者の実際の回答では活用されている場合も，表中ではすべて辞書形で代表させる。また，誤答①はコロケーションとしては成立するが，試行の例文の文脈には合わないもの，誤答②はコロケーションとして成立しないものを表す。

表5 例文１回答：日本の文化は，古くから中国や韓国の影響（　　）＿＿＿＿。

	中級	上級
前：正答	受ける (11) ある (1) 与えられる (1)	受ける (18) 与えられる (4) 及ぶ (3) ある (1)
後：正答	受ける (17) ある (1) 現れる (1) 及ぶ (1)	受ける (16) 無視できない (4) 現れる (4) 及ぶ (3) 与えられる (2) 及ぼされる (2) 他
前：誤答①	与える (1) 残る (1) 及ぼす (1)	与える (4) 及ぼす (4) もたらす (1)
後：誤答①	与える (13) 及ぼす (12) 生じる (1) 他	与える (4) 及ぼす (4) 残る (4) 出る (3) もたらす (2) 持つ (2) 他
前：誤答②	もらう (1) とる (1) 表わす (1) 関係する (1) 他	発展する (2) もらう (1) 得る (1) 繋がる (1) 寄せる (1) かける (1) 他
後：誤答②	とる (2) 含む (1)	異なる (1)

表５の例文１では，「かりん」を使用する前から「受ける」が両レベルで正答として多数選ばれている一方で，両レベルとも「かりん」使用後もヴォイス[6]による誤答は多く見られた。

また「かりん」使用後，特に上級では「無視できない」「現れる」など新たな意味の語彙も選択されている。

表６の例文２の中級では，「かりん」使用前は「する」が圧倒的であったが，使用後は「行う」「推進する」など「する」の類義語のほか，「着手する」「努める」といった異なる意味の語も多くが選択されており，上級でも同様の傾向が見られる。

一方で「誤答①」の「開発に従事する / 夢中になる / 励む / 尽力する」などは，開発を行う主体が「人」である時は使用可能であるが，例文２の「自動車会社」のような「組織」が主体の場合では多くの文脈で不自然さを感じると思われる。こうした主体に制限のある動詞についての解説は，一般的な辞書や本システムでも解説はなく，誤答の原因になり得ると思われる。

表6 例文２回答：ほとんどの自動車会社では，自動運転技術の開発（　　）　が，まだ，100%自動運転の車はない。

	中級	上級
前：正答	する (9) 成功する (2) 進む (1) 進行する 始める。 行う	進む (8) 取り組む (8) 行う (3) 注目する (2) 推進する (2) 始める (2) 他
後：正答	進める (18) 着手する (9) 進む (7) 推進する (7) 取り組む (5) 努める (3) 行う (2) 他	着手する (17) 進める (13) 推進する (7) 取り組む (7) 進む (6) 行う (6) 努める (3) 手がける (2) 他
前：誤答①	できる (1) 従事する (1)	夢中になる (1) 励む (1) 尽力する (1) 図る (1)
後：誤答①	促す (1)	促す (1)
前：誤答②	頑張る (2) 完成する (2) 研究する (2) 導入する (1) 起こす (1) 他	欠ける (1)
後：誤答②	（なし）	注ぐ (1)

次に表７の例文３は，「期待」と「不安」二つの名詞に続き，文脈に合う動詞としては「抱く」のみを正解とした。この例文では例えば中級の誤答①，上級の誤答②のように，「かりん」使用前に書いた動詞を使用後に記入せず，「かりん」で知った新しい動詞を書き，誤答になるケースが多く見られた。例えば「期待 / 不安」が「克服する / 募る / 寄せ

表7 例文 3 回答：私は今年の10月，胸に期待と不安（　　）　　て，日本に留学してきた。

	中級	上級
前：正答	抱く (3)	抱く (8)
後：正答	抱く (19)	抱く (17)
前：誤答①	持つ (12) 感じる (5)	持つ (9) 抱える (5) 感じる (1) 襲う (1) 駆られる (1)
後：誤答①	抱える (3) 克服する (1) 募る (1) なる (1) 覚える (1)	感じる (8) 抱える (7) 覚える (1)
前：誤答②	なる (3) する (1) 起こる (1) 思う (1)	膨らむ (2) 溢れる (1) 思う (1) ある (1)
後：誤答②	0	寄せる (2) 込める (1) 背負う (1)

る」は「かりん」使用後に使用されたが，「期待 / 不安」を「抱いて～する」が持つ「保持して～する」という意味を持たない。「込める / 背負う」は，「不安」とは共起をせず，また「込めて～」は単に「保持して～」以上の，「背負う」は「他者の～を保持して」という意味を持つ。このような動詞の個々の意味や特徴を知らずに適切なコロケーションを選ぶことは困難であり誤用の原因となっていると考えられる。

表 8 の例文 4 では，両レベルともヴォイスの間違いによる誤答が多かった。文の意味では被動者の動詞ではなく動作者の動詞を用いるべきであるのに，上級レベルにおいても，被動者の動詞（「被る」「受ける」）や自動詞（「生じる」「発生する」）などが選択され，回答数も動詞の種類も増えている。これは不注意によるものか「被害」という受動的意味を表す名詞に引きずられたものかなど原因は不明であるが，ヴォイスについては上級レベル

表8 例文 4 回答：先月の台風は，西日本に大きな被害（　　）　　。

	中級	上級
前：正答	与える (2) 引き起こす (1)	与える (7) 引き起こす (6) もたらす (6) 及ぼす (5) 他
後：正答	もたらす (3) 与える (2) 及ぼす (2) 引き起こす (1)	もたらす (10) 与える (8) 及ぼす (5) 引き起こす (4) 発生させる (1) 他
前：誤答①	受ける (8) ある (6) 被る (4) 遭う (4) 発生する (4)	受ける (3) 及ぶ (3) 被る 遭う (2) 他
後：誤答①	受ける (16) 被る (12) 遭う (5) 発生する (2) 他	被る (8) 受ける (7) 生じる (5) 遭う (4) 及ぶ (2) 発生する (1) 起こる (1) 出る (1) 他
前：誤答②	する (3) できる (1) 起きる (1) 起こす (1) 作る (1)	あげる (1)
後：誤答②	もらう (1) できる (1) 防ぐ (1) 軽減する (1)	

でも選択が難しいという点は，特記すべきと考えられる。

表 9 の例文 5 では，両レベルとも，「かりん」使用前から「始める」「する」など動作の始動や実行を表す動詞が選ばれてきたが，「かりん」使用語は，同様の意味の類義語も多く選択されている。

また誤答を見ると，特に中級の「かりん」使用前の誤答では，「書く / 提出する / 勉強する」といった動詞を回答した学生が見られるが，これは中国語では「作業」に「宿題」という意味があるためと考えられる。このように特に中級レベルでは，N2，N1の名詞や動詞についての知識が不足している場合があり，

表9 例文5回答：私の説明が終わったら，作業（　　）　　ください。

	中級	上級
前：正答	する (6) 始める (5) 続ける (2) 他	始める (14) する (8) 開始する (3) 続ける (3) 他
後：正答	進める (14) 開始する (6) 行う (4) 他	開始する 始める 行う 進める
前：誤答①	やめる (1)	取り組む (1) やめる (1) 入る (1)
後：誤答①	従事する (4) 終える (3) 繰り返す (2) 他	終える 終わる 進行する 集中する他
前：誤答②	出す (4) 書く (2) 作る (2) 始まる (2) 提出する (1) 勉強する (1) 他	続く (1) 止める (1) 展開する (1) 考える (1)
後：誤答②	とる (3) できる (1) 直す (1) 勉強する (1)	0

誤答の原因のひとつになっていると考えられる。

5.2.2　試行結果の考察

このように，全般的に両レベルとも「かりん」を使用することによって，学習者がこれまで通常作文の際に行ってきた方法より回答数も正答数も増え，文脈に合わせてより多様な語の選択が行われているという結果を得た（今回の試行の目的の①②）。このことから「かりん」については，中級レベルから十分使用可能であることがわかり，コロケーションの選択に有益であることも確認された。

一方で，それぞれの例文では上述のような誤答も見つかっており，学習者がより適切なコロケーションを選択するためには，以下のような情報を加えていくことが重要と思われる。

ひとつめは，例文2や4に見られたようなヴォイスの問題である。この誤答は上級レベルでも多く見られることから，大変重要な課題と考えられる。改善点としては，このようなヴォイス的な意味を持つ動詞については，現在のような例文ばかりでなく，文型がわかるような単純化された例文（例：台風が山口に被害を与えた）などを加え，文型をよりわかりやすくさせ，意識を向けさせるという方法が考えられる。

ふたつめは，例文5の「作業」のように母語（特に漢語を用いる中国語）との意味の違いが見られる語についての問題である。例えば「作業」では，「○作業を行ってください。×作業を出してください。「作業」は「製作や操作などの仕事をするという動作を表す名詞で，『宿題』という意味はありません」と言った具体的な説明を加えることが単純で効果が高いと考えられる。

一方，例文2で見られたような動作の主体が「人」か「組織」かによって使用が難しいなどの動詞の使用における制限などについては，情報は必要であるが，記述は難しいことが予想される。これは現存する日本語の大規模コーパスでは，例えば主体と動作を結び付けた検索は難しく，記述を加える動詞の制限の正確さも直観に頼ったものとならざるを得ないためである。この点については，先行研究を検討した上で，今後の課題としたい。

5.2.3　アンケート結果と考察

以下，上記の試行を行った後に実施したアンケートの結果について述べる。調査では，(1「かりん」は役に立ったか，2(1)画面（見やすさ・美しさ），(2) 操作性（使いやすさ・わかりやすさ），(3) 例文 (4) 翻訳（長さ・わかりやすさ）について5段階評価で回答，さらにその理由を自由に記述してもらった。

以下，これらの質問に対する5段階評価を図11に示し，設問ごとの理由（自由記述）

」

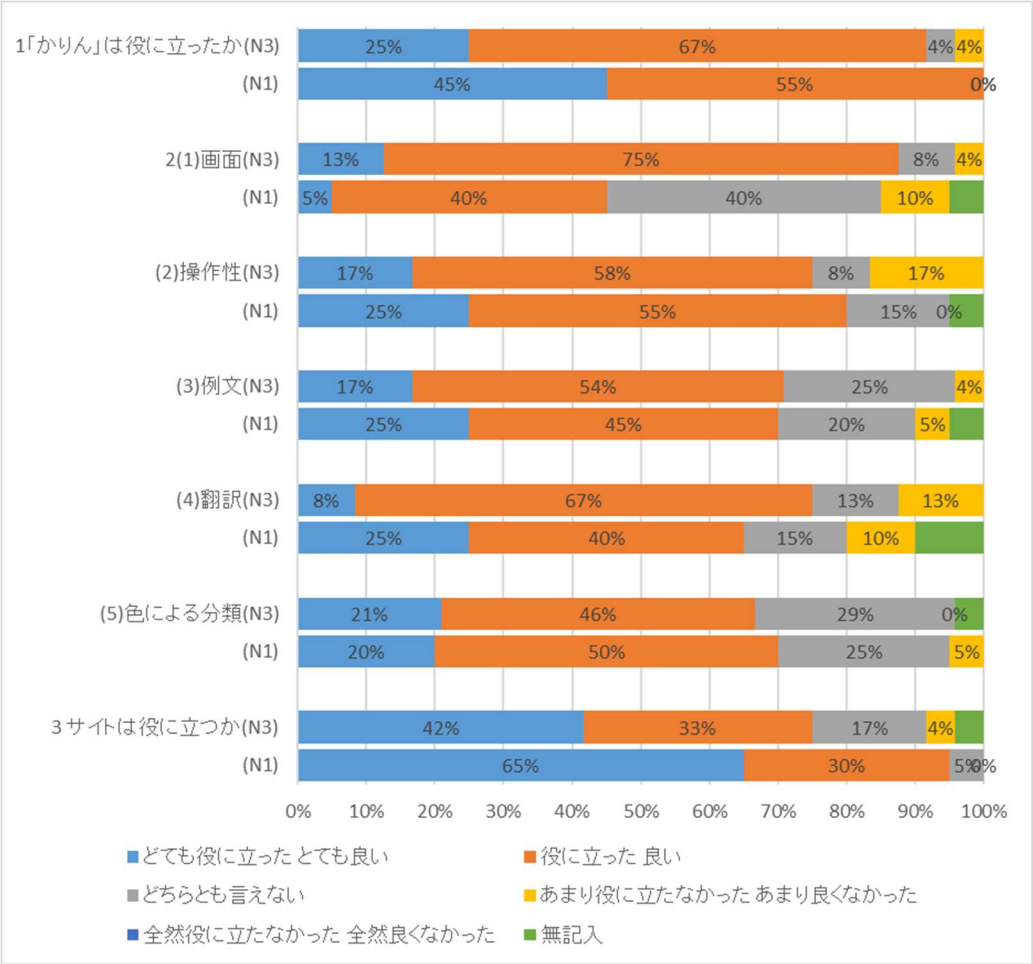

図11　「かりん」アンケート結果

の一部を記す。またコメント末の（　）内の最初の文字はレベルを（中・上），「／」の右は学習者の母語の先頭文字を示し，「ア他は中国語・韓国以外のアジア言語，「ヨ他」英語以外のヨーロッパ言語を表す。

結果，すべての項目について，5段階評価では「とても役に立った／よい」，「役に立った／よい」を加えると過半数を超えていた。中でも「かりん」が（試行の回答に）「役に立ったか」では，「とても役に立った」「役に立った」を合計すると，両レベルとも90%を超えている。コメントでは，中級レベルは「新しいことばを覚えるとき，発音と書き方を覚えるだけで，文の中での使い方や名詞と動詞の組み合わせの仕方などはわからないから」（中／中），「名詞を使うとき，たぶんどの動詞を使うかを知るのは難しい。GUESSしなければいけない。かりんはこのような問題の支援をしてくれる」（中／ア他）など名詞の使い方の役に立つといった記述が目立ったのに対し，上級では「より文章の表現力を豊かにすることができます。また，自然な日本語を話すことに役立ちます」（上／中），「大きくバリエーションが書けなかったので，かりんで検索したら色々な助詞と動詞があり，役に立ったと思いました」（上／ヨ他）など，自然な日本語や表現の豊かさに言及するコメントが目立った。

また最後のこの「サイトは役に立つか」という質問には，特に上級レベルの学生の評価が高かった。理由として「レポートを書くときに役立つ」（上／韓），「日本語の文を書くとき，自然なコロケーションがわかるために，辞書をひくことよりもっと便利でちゃんとした例もあるので，とても役に立った」（上／中）など，「かりん」が書き言葉のコーパスを基にしているため，レポートなどを日本語で書く機会がある上級の学習者がコロケーションの必要性を実感し，本システムの有効性を感じたことが窺える。

また，２の学習者画面，学習者のための支援機能については，ほぼレベルの差がなかったものの，(2)「画面」では上級レベルの満足度が低い。これは「色の組み合わせはもう少し考えるほうがいいと思います。明るすぎかもしれません。デザインも同じです」（上／中）のように，ほとんどが配色やデザインに関するものであったが，「簡単でいいと思う」（上／中）という意見がある一方で「シンプルすぎ」（上／韓）のように機能も関わると思われる意見も散見された。

また(4)「翻訳」では中級の評価が上級を若干上回っているのは，より翻訳への依存度が高いためと思われる。しかし英語の翻訳が他の二言語に比べ，極端に少ないことから「ヨーロッパ言語がないので，理解できない」（中／ヨ他），「英語の意味があれば，もっといい」（中／ア他）のように，中国語圏，韓国語圏以外の学習者の共通言語としての英語の整備が急務であることが改めて確認された。中級レベルからのすべての学習者に有益なシステムとなるよう努力をしていきたい。

6.　おわりに―今後の課題

以上，日本語中上級学習者のためのコロケーション検索システム「かりん」についての特徴を操作方法に沿って紹介するとともに，2019年度に実施した試行・アンケート結果を基に，今後の課題について考察した。

本システムは，従来のコロケーション検索システムに比べ，使用できる学習者のレベルを広げたことから，日本語レベルによっては学習者がそれまで知らなかった動詞を含むコロケーションも選択することが可能になる。そのため名詞や動詞について意味や用法の丁寧な解説がないと誤用を増やす可能性があることが今回の試行を通じて大きな問題であることがわかった。今後は，今回の結果を参考にどのような情報を優先的に加えるべきか改めて検討し，さらに改善を行っていきたい。

「日本語中上級学習者のためのコロケーション検索システム《かりん》URL」
http://Japanese-learning.isc.yamaguchi-u.ac.jp/collocation/

（留学生センター　教授）
（大分大学国際教育研究センター　教授）
（大分大学国際教育研究センター　講師）

【謝辞】

本研究は科研費（基盤研究（C)16K02813）の助成を受けたものです。

【参考文献】

石川慎一郎（2008）『英語コーパスと言語教育』，大修館書店

国立国語研究所コーパス開発センター，2011，『日本語書き言葉均衡コーパス』検索アプリケーション「中納言」
https://chunagon.ninjal.ac.jp/auth/login

徳弘康代，2006，日本語教育における中上級漢字語彙教育の研究”，早稲田大学大学院日本語教育研究科　博士論文
http://dspace.wul.

waseda.ac.jp/dspace/bitstream/2065/5428/1/Honbun-4252.pdf
村木新次郎,1991,『日本語動詞の諸相』ひつじ書房

【注】

1) 国立国語研究所・Lago言語研究所：http://nlb.ninjal.ac.jp/
2) 筑波大学・Lago言語研究所：”Web コーパス” http://nlt.tsukuba.lagoinst.info/
3) 日本語学習支援システム　日本語共起語検索システム「なつめ」 https://hinoki-project.org/natsume/
4) 石川（2008）は（コロケーションの共起強度を考える際に）共起頻度だけでは中心語と共起語の語自体の頻度によって値が大きく左右されると述べ，また共起強度を示す指標には高頻度のコロケーションの評価に強い頻度型指標（Ｔスコア，対数尤度比）や低頻度でも特徴的なコロケーションを検出する非頻度型指標（相互情報量）があるが，Dice係数はその中間型であるとしている。こうした理由から本稿ではDice係数を学習者に提示する指標として採用した。
5) 本稿では，「正答」は文脈に沿った意味の動詞を選んでいる場合とした。したがって助詞の間違いやテンスやアスペクト，活用形などの文法的な間違いがあっても選んだ動詞が正しければ正答とした。ただし，例文１のように「受ける」を選ぶべきところを「与えられる」のようにヴォイスを変えることで文脈に合わせた場合は正答とした。
6) 「与える／受ける」のような動詞の対立について村木(1991)は，「太郎は次郎に注意をあたえた／次郎は太郎から注意をうけた」の例を挙げ，ともに「太郎」は動作者で「次郎」が被動者であり，二つの関与者は対応する二つの文で格が交替していることや，これらの述語部分は「注意した／注意された」とそれぞれ交替することから，このような語彙によって「能動／受動」の対立をつくる場合を「迂言的な受動表現」と呼び，ヴォイスの一部を成すと考えている。本稿でも語彙による「能動／受動」の対立も，以下「ヴォイス」と呼ぶこととする。