コーパス検索アプリケーション「中納言」を利用したカラ格受動文の調査

和氣　愛仁

キーワード：コーパス、形態素解析、カラ格、受動文

要　旨

「現代日本語書き言葉均衡コーパス（BCCWJ）」の検索アプリケーションとして公開されている「中納言」を用い、カラ格受動文の用例検索を行う。本来形態論的情報しか付与されていない「中納言」であるが、一定の検索条件を与えることによって、構文論的な条件による用例検索が擬似的に可能な場合があることを示す。また実際に検索されたカラ格受動文の用例を分析することにより、これまで主として研究者の内省によってきた言語研究にさらなる信頼性を与えるために、大規模な均衡コーパスによる用例検索が有効であることを示す。

1. 本稿の目的

本稿では、コーパス検索アプリケーション「中納言」を利用してカラ格受動文の用例を検索し、先行研究により指摘されたカラ格受動文に関する知見の検証と、未発見の事実の探索およびその原理の解明を最終的な目的としつつ、その第一段階として、実践的な用例検索技術および実際に検索された用例の分析について述べる。以下、「中納言」を用いて、基本的には形態論情報による検索を行いつつ、検索条件を工夫することによって、意味・構文論的な分析までを視野に入れた用例検索を試みる。本稿の前半では「中納言」の紹介をしつつ用例検索の実際について述べ、後半で実際に検索された用例について若干の分析を行う。

2. 「中納言」の基礎技術

2.1. 「中納言」について

ひとくちにコーパス検索アプリケーションと言っても、テキストの形態素解析を

目的とするChaSen・MeCabや、係り受け解析を目的とするCaboChaのような、どちらかといえば自然言語処理を目的として汎用的に基礎データの解析を行うようなタイプのアプリケーションから、XML等により一定のルールに則ってタグ付けされたテキストデータの中から目的とする構造を持つテキストを抽出するChaKi・ひまわりのような、どちらかといえば言語研究を目的として利用されるようなタイプのものまで、様々なものが開発されてきている。

2011年8月に公開された「中納言」は、「現代日本語書き言葉均衡コーパス(BCCWJ)」の検索ツールとしてウェブ上で運用されているコーパス検索アプリケーションである。BCCWJは人間文化研究機構国立国語研究所と文部科学省科学研究費特定領域研究「日本語コーパス」プロジェクトが共同で開発した均衡コーパスであり、書籍・新聞・雑誌・白書・ウェブ掲示板やブログなどの多様なソースから、2011年11月末時点で約1億語が収録されている。形態素解析ずみのテキストデータには、形態論情報のほか文章構造や書誌情報なども付与されており、詳細な用例検索が可能となっている。BCCWJそれ自体、例えば言語情報処理技術者が言語データ解析のために利用したいといった場合には、一定の手続きに従ってデータ全体の利用許諾を得ることができるが、言語研究者のように、自力でデータ解析を行うまでのことは必要ないが、解析済みのデータから必要なデータを抽出したいといった目的を持つ場合には、より限定的な機能を持つアプリケーションとして、「少納言」「中納言」というふたつのアプリケーションがウェブ上で利用できる。このうち「少納言」は、フレーズ検索のみが可能なもの(形態論情報による検索は不可)で、利用許諾契約の締結は不要(無償で利用可能)、検索できる用例数は最大500、前後文脈が正規表現により指定可能、となっている。一方の「中納言」は、形態論情報によるより詳細な検索が可能で、利用許諾契約の締結が必要(利用料年額3150円、ただし当面の間は無償)、検索できる用例数は無制限かつダウンロード可能、前後文脈の条件を最大10まで指定可能、となっている。詳細は後述するが、言語研究者ならば、是非とも利用許諾を得た上で、「中納言」を利用していただきたい。「中納言」は、単なるフレーズ検索を越えた検索条件指定の自由度の高さが非常に魅力的であり、これまでコーパス言語学に興味のなかった研究者や用例の収集に苦慮していた研究者にとっても、興味深い用例を手軽に数多く抽出できる能力を秘めているという点で、非常に強力な研究用の基礎ツールとなり得る。

ただ、残念ながら、2011年11月末現在、「中納言」をホストしているサーバの能力が必ずしも十分ではないようで、前述の文脈指定条件数内であっても、やや複雑な検索条件を与えた場合や、大量の検索結果が得られた場合には、サーバからの

応答が返ってこなくなることがあった。また、検索はなんとか可能であっても、ダウンロードは不可能ということもあった。以下随時その旨補足してゆくが、そのような場合でも、検索条件を工夫することにより、目的とする結果と近い結果を得ることが可能なことがある。

なお、サーバの能力については、担当者に問い合わせたところ、2011 年度中の増強を検討しているとのことであった。今後、より複雑な検索条件指定、より大量の検索結果取得が可能になることを期待している。

2.2. 形態素解析と「短単位」「長単位」

「短単位」「長単位」とは自然言語処理の形態素解析における技術用語であり、一般の言語研究者には聞き慣れない用語かもしれない。日本語は分かち書きをしない言語であり、計算機上で日本語を扱う場合、英語などとは異なり、そもそも語の区切りがどこであるのかということ自体が極めて大きな問題となる。そこで用いられるのが、形態素解析技術である。形態素解析は、平たく言えば、ひと続きの文字列から、自動的に「語」を切り出す技術と言ってよいだろう。その時に利用されるソフトウェアが ChaSen や MeCab 等の形態素解析アプリケーションであり、それと組み合わせて使われるのが IPADIC や UniDic 等の形態素解析用辞書である。

以下に、MeCab（辞書は naist-jdic）を用いて日本語文の形態素解析を行った例を示す。１行目がコマンド入力、２行目が解析対象とした日本語文、３行目以降が実際の解析結果である。解析結果の表示は、各語について左から、表層形、品詞、品詞細分類 1、品詞細分類 2、品詞細分類 3、活用形、活用型、原形、読み、発音、表記ゆれ情報、複合語情報、となっている[*1]。

*1 naist-jdic は奈良先端科学技術大学院大学により作成・公開されている形態素解析用辞書で、IPADIC をもとに、表記ゆれ情報、複合語情報を追加したものである。(1) では、naist-jdic を用いたため、MeCab デフォルトの表示項目に加え、表記ゆれ情報、複合語情報が表示されている。

(1) MeCab による形態素解析の例

```
[wakit@ict ~]$ mecab
太郎が次郎に英語を教えてやった。
太郎　名詞,固有名詞,人名,名,*,*,太郎,タロウ,タロー,,
が　助詞,格助詞,一般,*,*,*,が,ガ,ガ,,
次郎　名詞,固有名詞,人名,名,*,*,次郎,ジロウ,ジロー,,
に　助詞,格助詞,一般,*,*,*,に,ニ,ニ,,
英語　名詞,一般,*,*,*,*,英語,エイゴ,エイゴ,,
を　助詞,格助詞,一般,*,*,*,を,ヲ,ヲ,,
教え　動詞,自立,*,*,一段,連用形,教える,オシエ,オシエ,おしえ/教え,
て　助詞,接続助詞,*,*,*,*,て,テ,テ,,
やっ　動詞,非自立,*,*,五段・ラ行,連用タ接続,やる,ヤッ,ヤッ,,
た　助動詞,*,*,*,特殊・タ,基本形,た,タ,タ,,
。　記号,句点,*,*,*,*,。,。,。,,
EOS
```

ただし、「語」を切り出すと言っても、そもそも「語」とは何かという問題がある。「語」の定義自体、実用的な技術開発を目指す自然言語処理と、学術的な研究理論の構築を目指す言語学とでは、自ずと異なってくるのが当然であり、形態素解析により分析される「語」は、言語学的な意味で言う語とは異なる単位であることもあり得る。また同じ自然言語処理であっても、目的とする処理によって、より短い単位でテキストを切り出した方が都合がよい場合もあれば、より長い単位で切り出した方が都合がよいこともある。そのようなことから、形態素解析においては、目的に応じてテキスト抽出の粒度を調整することが行われる。この短い単位で切り出した「語」を短単位と呼び、長い単位で切り出した「語」を長単位と呼んでいる。言語研究に形態素解析を利用する立場から言うと、語彙論あるいは形態論的な研究には短単位が、構文論的な研究には長単位が向いているとは言えるが、研究上の手法によっては常にそれが当てはまるわけでもない。

以下はそれぞれ、「中納言」を用いて、[朝食] [を] [食べる] という語彙素の連続を含む用例を、短単位および長単位により検索した例である。[食べる] をキー（中心となる検索語）とし、その前後文脈をともに 30 語表示してある。語の区切りは "|" による。

(2) 「中納言」による形態素解析の例（キーをゴシック体で表示）

a. 短単位解析

を|検査|する|こと|で|小児|生活|習慣|病|の|予防|、|早期|発見|や|治療|に|取り組ん|で|い|ます|。| |先月|号|で|は|、|(朝食)|(を)|**食べる**|こと|は|脳|の|働き|を|活発|に|し|、|一|日|の|活動|を|活性|化|し|て|くれる|大事|な|習慣|で|ある|こと|を|紹介|し|

b. 長単位解析

コレステロール値|、|肥満度|の|３つ|の|項目|を|検査する|こと|で|小児生活習慣病|の|予防|、|早期発見|や|治療|に|取り組ん|でい|ます|。| |先月号|で|は|、|(朝食)|(を)|**食べる**|こと|は|脳|の|働き|を|活発|に|し|、|一日|の|活動|を|活性化し|てくれる|大事|な|習慣|である|こと|を|紹介し|まし|た|。|反対|に|朝食|

（『広報ひゅうが』宮崎県日向市 2008）

「小児生活習慣病」「早期発見」「ている」「てくれる」「である」等の分析結果が異なることが見て取れるであろう。

現在では形態素解析技術は相当程度進歩しており、例えば上述の MeCab を用いた際の解析精度は、MeCab 開発者によれば 98%程度とされている。ただし、解析結果については、人手による辞書の調整が関与するとはいえ、基本的には計算機が自動的に行うことであるので、当然のことながら完全に誤りがないというわけではない。こうした自然言語処理に基づいた言語研究を行う場合は、そのことを常に意識しておく必要がある。

なお、「中納言」における短単位解析には MeCab と UniDic が、長単位解析には Uchimoto & Isahara (2007) による長単位解析器が、それぞれ用いられている（後者については国立国語研究所山崎誠氏の御教示による）。

3.「中納言」によるカラ格受動文の抽出

さて、ここから、実際に「中納言」を用いた用例検索について述べていく。「中納言」の利用方法については、サイトにもマニュアルが掲載されているのでそちらを参照していただきたい。ただし、本稿で触れる、検索条件式による検索および履歴による検索については記載がないので、本稿中で随時解説していく。

「中納言」により検索可能なのは、あくまでも表層形式として顕現した語の連接

であり、それぞれの語に付与されている情報も、あくまで形態論情報である。したがって、本稿が目的とするような、意味論・構文論的な情報抽出を目的とした用例検索は、「中納言」が本来目的とするものではない。しかしながら、特定の条件に当てはまるような文型の場合、検索条件を工夫すれば、目的とする用例を比較的手軽に抽出できる。本稿ではそのような例として、カラ格受動文を取り上げる。

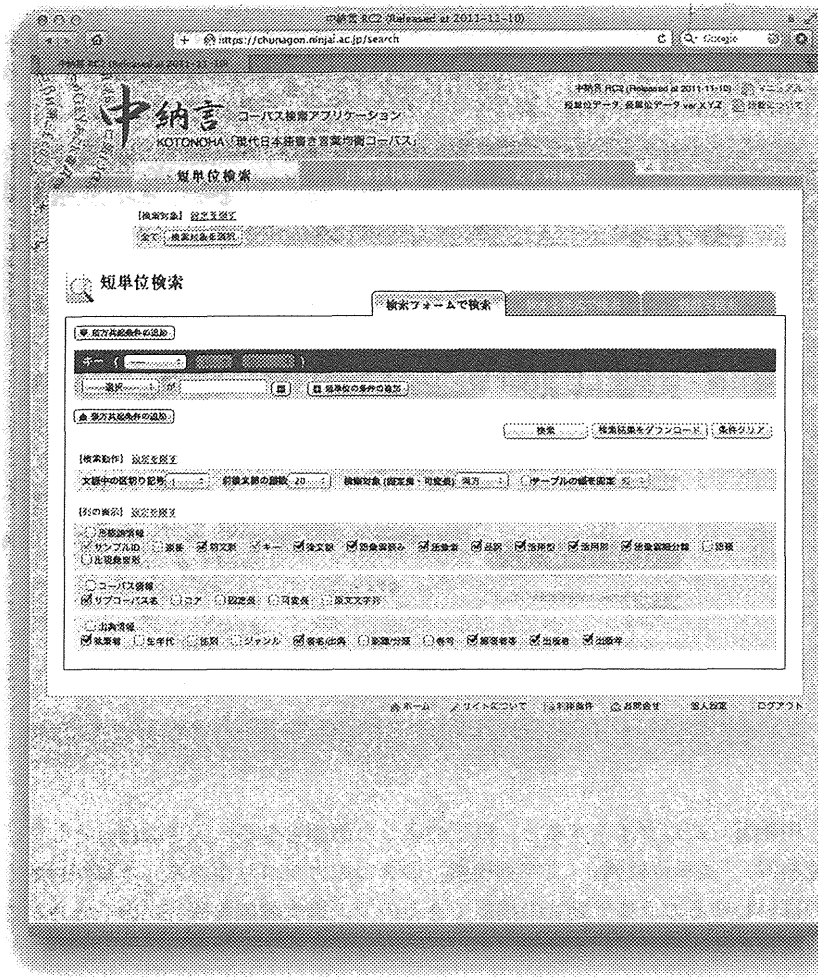
なお、当然のことながら、ここで検索された用例数や用例タイプごとの数的比率が、直ちに定量的な研究のためのデータとして利用できるわけではない。取り漏らした用例もあるだろうし、目的に合致しない不要データも多数入り込む余地がある。したがって、言語研究のための基礎データとするには、現時点ではどうしても人手による用例のチェック作業が欠かせないし、得られた用例の数や比率についても直ちに意味のあるものと見なすことはできない。将来的には、意味論・構文論的な研究において、計算機を用いた自動的かつ定量的な研究が可能になる日も来るかもしれないが、残念ながら現在はまだそこまでは至っていない。現時点では、特に一般的な言語研究においては、「中納言」は当面、データ収集ツールとしての側面が強くクローズアップされることになるだろう。

3.1. ログイン

以下は「中納言」のログイン画面である。利用許諾契約により発行されたユーザ ID およびパスワードを入力し、システムにログインする。



「中納言」にログインすると、以下のような画面となる。大きな区分として「短単位検索」「長単位検索」「文字列検索」の３つのカテゴリがある。本稿が目的とするのはいわゆるカラ格受動文の抽出であるので、構文論的な研究に向くとされる長単位検索を第一選択として検索を実行してみる。



3.2. 検索条件の指定

ここではまず、[ヒト性名詞] [カラ] [動詞] [レル/ラレル]というパターンを抽出してみる。動詞をキーとし、前方共起条件として、キーの１語前に格助詞カラ、キーの２語前にヒト性名詞、また後方共起条件として、キーの１語後ろに助動詞レルまたはラレルが共起する、というパターンである。まず[長単位検索]タブをクリックし、[検索フォームで検索]タブの中に、以下の手順で検索条件を入力していく。

最初に、[キー]とある行の下、[---選択---]ドロップダウンリストから[品詞]を選択し、[大分類]が[動詞]となるように設定する。これにより、キーとして動詞を抽出することが指定される。キーとして指定した語は、読み、原文中での表記、活用型、原文中での活用形など、分析に有効な情報が出力される。ここでは、後に動詞のリストを作ること等もふまえ、動詞をキーとしておく。



次に、[後方共起条件の追加]ボタンをクリックし、[後方共起 1]とある行の下、[---選択---]ドロップダウンリストから[WHERE 句]を選択し、右隣のテキストボックスに「語彙素 = "れる" OR 語彙素 = "られる"」と入力する（カギ括弧は除く。以降同様）。[---選択---]ドロップダウンリストの内容は右の通りである。

ここではつまり、動詞に後続する要素として、「れる」または「られる」という基本形式を持つ語を指定したいわけだが、複数の語彙素の条件をそれぞれ指定しようと[長単位の条件の追加]ボタンをクリックしても、AND 条件でしか条件を追加することができない。OR 条件を指定したい場合は、[---選択---]ドロップダウンリストから[WHERE 句]を選択し、条件の内容として前述のような式を入力することで、希望の動作を実現できる。なお、ひとつのテキストボックスに入力された条件式は、自動的に全体を括弧()でくくられたものとして計算されるので、（通常 AND が OR に優先して評価されるのと異なり）テキストボックス内の OR 条件が、複数の共起条件間の AND 条件よりも優先して評価される。

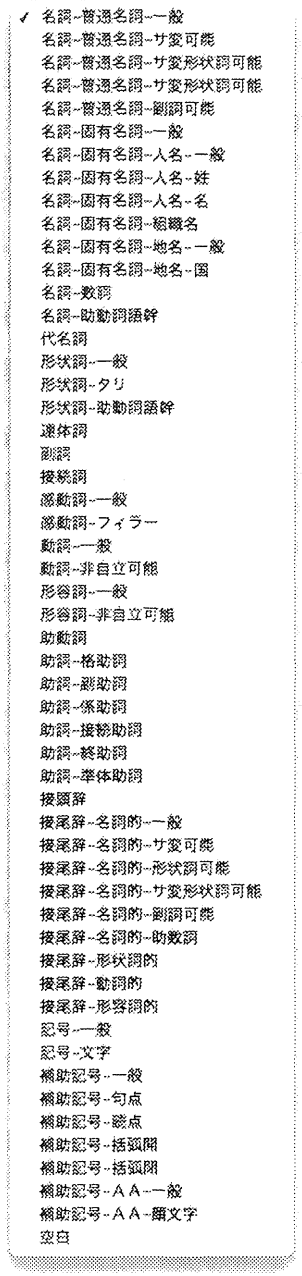
次に、[前方共起条件の追加]ボタンをクリックし、[前方共起 1]とある行の下、[---選択---]ドロップダウンリストから[語彙素]を選択し、右隣のテキストボックスに「から」と入力する。さらに右隣の[長単位の条件の追加]ボタンをクリックすると、"AND"で始まる行が追加されるので、さらに２行目の[---選択---]ドロップダウンリストから[品詞]を選択し、[小分類]が[助詞-格助詞]となるように設定する。これにより、前方共起条件１として、動詞のひとつ前の語が格助詞カラであることを指定したことになる。

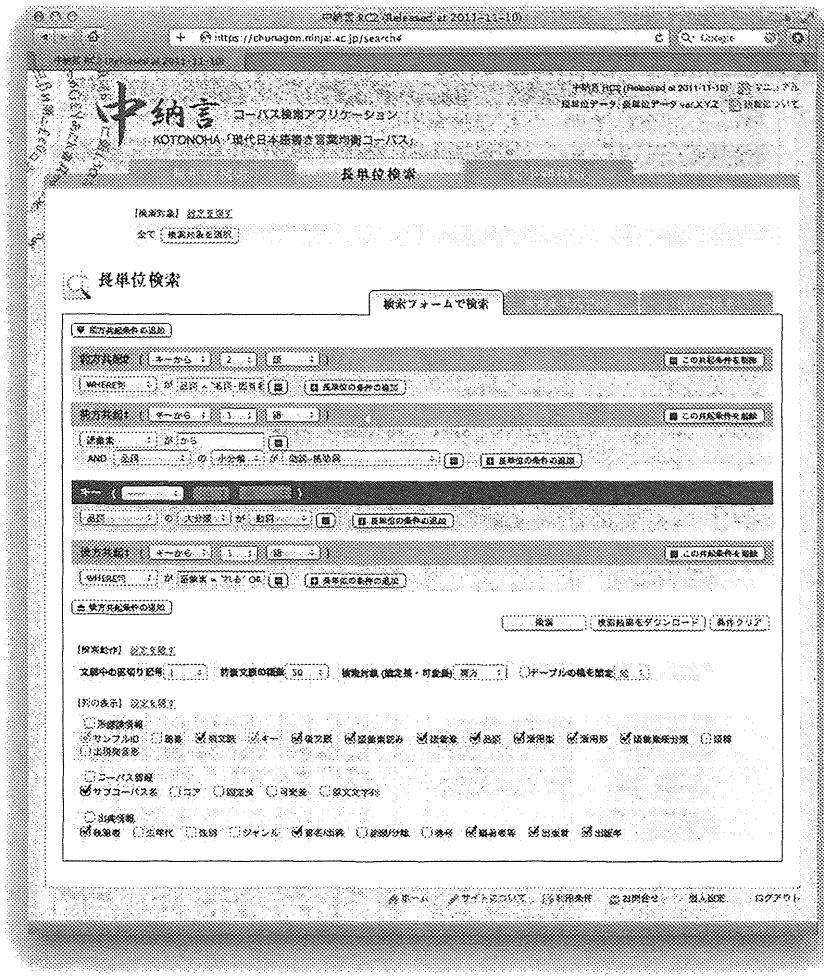
次に、[前方共起条件の追加]ボタンを再度クリックし、[前方共起 2]とある行の下、[---選択---]ドロップダウンリストから[WHERE 句]を選択し、右隣のテキストボックスに「品詞 = "名詞-固有名詞-一般" OR 品詞 LIKE "名詞-固有名詞-人名%"」と入力する。ここでの LIKE は、文字列比較演算子の一種で、ワイルドカード("%")を利用することができる。つまりここでの指示は、品詞が"名詞-固有名詞-一般"であるもの、または、品詞が"名詞-固有名詞-人名"で始まるもの、という意味になる。

「中納言」の品詞情報として用意されているのは右の通りである。

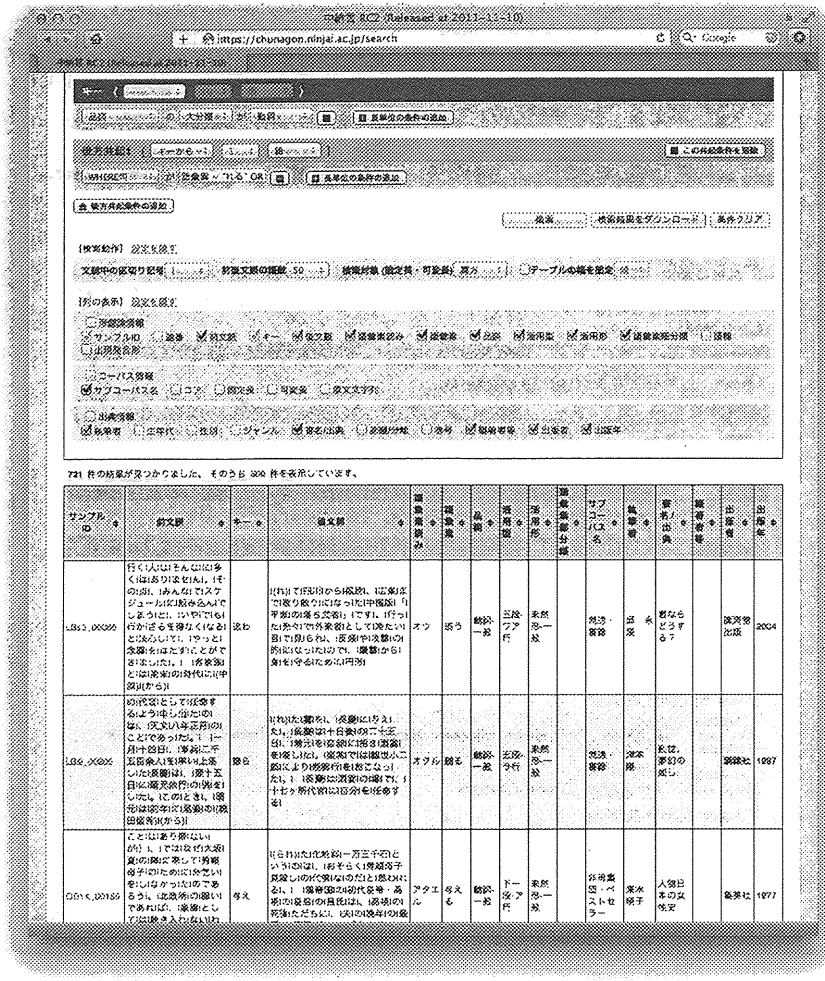
つまり、「品詞 LIKE "名詞-固有名詞-人名%"」という条件式は、上のうち、"名詞-固有名詞-人名-一般"、"名詞-固有名詞-人名-姓"、"名詞-固有名詞-人名-名"にマッチすることになる。以上をまとめて、前方共起条件２では、動詞のふたつ前の要素が、名詞であり、かつその内容が、固有名詞一般もしくは固有名詞のうち人名であるもの、という条件を指定したことになる。つまりここでは、「ヒト性名詞」に近似する条件として、固有名詞のうちヒトらしいものを条件として指定した、ということである。より狭い範囲で名詞の条件を指定したければ、「品詞 = "名詞-固有名詞-一般" OR 」の部分を消去してもよいだろう。

ここまでの検索条件を画面上で指定すると、以下のようになる。





これで[検索]ボタンをクリックすれば、以下のように検索結果が表示されるはずである。ただし、前述の通り、現状の「中納言」は、このくらいの複雑さの検索条件を指定するとかなり反応が遅くなる。筆者が試したところ、上記条件で検索結果と検索件数が表示されるまでに、約３分半ほどかかっている。



2011 年 11 月末現在、この条件での検索件数は 721 件であった。もちろん、前述の通りすべての例がいわゆるカラ格受動文であるわけではなく、最終的には人手で用例を検証する必要がある。

その他、検索のオプションとして、前後文脈中の語の区切り記号、前後文脈の語数、検索対象データ等の指定ができる。また、検索結果の表示オプションとして、形態論情報、コーパス情報、出典情報の表示・非表示を切り替えることもできる。これらについての詳細は、「中納言」のオンラインマニュアルをご参照いただきた

い。

また、[検索結果をダウンロード]ボタンを押せば、検索結果すべてがタブ区切り形式テキストファイルでダウンロードできる。ただしこれも上述のサーバ性能のせいか、数千件を超える検索結果の場合に反応が返ってこなくなることがあった。これについてもぜひ改善を期待したいところである。

3.3. 検索履歴の再利用

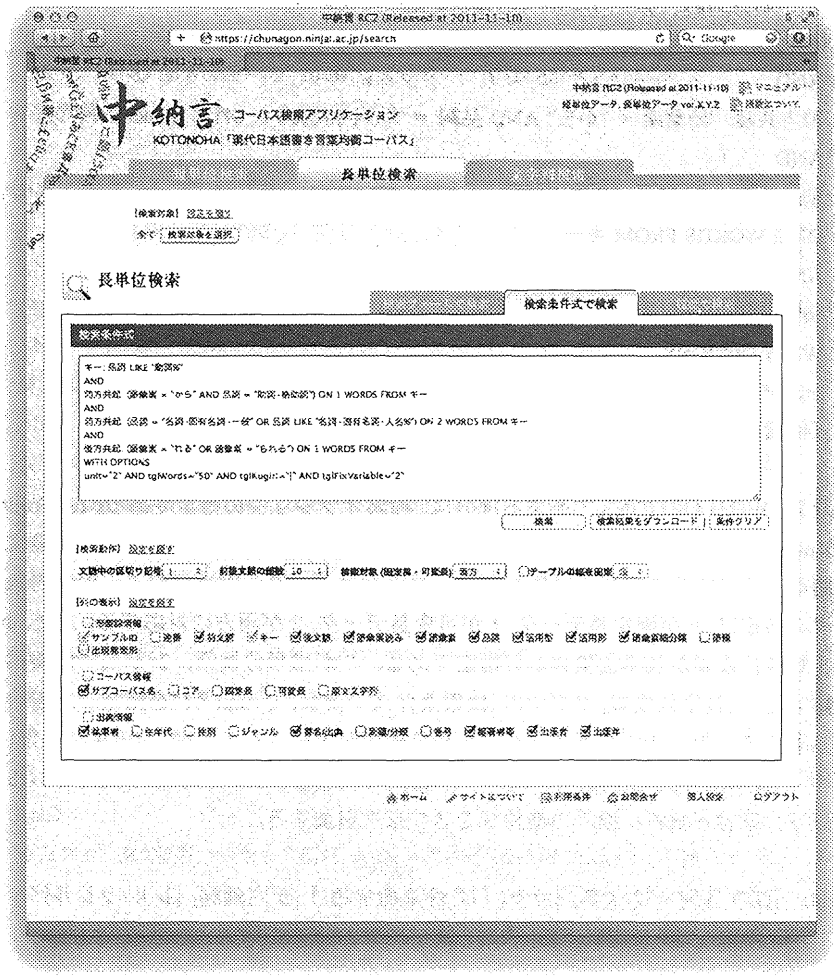
検索フォームで一度検索を行うと、その時に指定した検索条件が、[履歴で検索]タブの中に保存されていく。画面例を以下に示す。



検索フォームで毎回検索条件を設定するのは少々面倒なので、以前と同じ条件で再度検索したい場合や、以前の検索条件の一部を変更して検索したい場合には、これを利用するのが便利である。同じ条件を利用したい場合は、該当の検索条件を選択して[検索]ボタンをクリックする。すると、検索フォームで検索したときと同様

に、検索結果が表示される。

以前指定した検索条件の一部を変更して検索したい場合は、もととなる検索条件を選択した上で、[編集して検索]ボタンをクリックする。すると、[検索条件式で検索]タブに移動し、選択した検索条件式がテキストエリアに表示されるので、ここで条件式の編集を行う。画面例を以下に示す。ここではわかりやすいように検索条件式に適宜改行を入れてある。



ここに表示されているのは、前述した検索フォームによる検索で使用された条件式、すなわち、[固有名詞一般/人名] [カラ] [動詞] [レル/ラレル] というパターンを抽出する条件式である。SQL になじみのある者ならわかりやすいかもしれないし、そうでなくてもある程度意味は推測できるであろう。検索フォームでの条件指定がどのような式として表現されているか、よく見比べていただきたい。

(3) [固有名詞一般/人名] [カラ] [動詞] [レル/ラレル]

キー: 品詞 LIKE "動詞%"
AND
前方共起: (語彙素 = "から" AND 品詞 = "助詞-格助詞") ON 1 WORDS FROM キー
AND
前方共起: (品詞 = "名詞-固有名詞-一般" OR 品詞 LIKE "名詞-固有名詞-人名%")
ON 2 WORDS FROM キー
AND
後方共起: (語彙素 = "れる" OR 語彙素 = "られる") ON 1 WORDS FROM キー
WITH OPTIONS
unit="2" AND tglWords="50" AND tglKugiri="|" AND tglFixVariable="2"
(検索結果：721 件)

なお、WITH OPTIONS で始まる部分は、検索オプションの指定である。tglWords、tglKugiri、tglFixVariable は、その名前から推測できるように、それぞれ、前後文脈の語数、前後文脈中の語の区切り記号、検索対象データの指定である。検索対象データの指定は、0 が固定長データ、1 が可変長データ、2 が両方に対応する。また、unit は、短単位検索か長単位検索かの指定である。1 が短単位検索、2 が長単位検索に対応する。[長単位検索] タブ中の [検索条件式で検索] で unit="1"を指定した場合、検索条件式による指定が優先され、短単位検索が実行される。逆に [短単位検索] タブ中の [検索条件式で検索] で unit="2"を指定した場合、長単位検索が実行される。

以下、筆者が試みた検索の条件式を３つほど追加する。

(4) [固有名詞一般/人名] [カラ] [(条件未指定語)] [ヲ] [動詞] [レル/ラレル]

キー: 品詞 LIKE "動詞%"
AND
前方共起: 語彙素 = "を" ON 1 WORDS FROM キー

AND
前方共起: (語彙素 = "から" AND 品詞 = "助詞-格助詞") ON 3 WORDS FROM キー
AND
前方共起: (品詞 = "名詞-固有名詞-一般" OR 品詞 LIKE "名詞-固有名詞-人名%")
ON 4 WORDS FROM キー
AND
後方共起: (語彙素 = "れる" OR 語彙素 = "られる") ON 1 WORDS FROM キー
WITH OPTIONS
unit="2" AND tglWords="50" AND tglKugiri="|" AND tglFixVariable="2"
(検索結果：170 件)

(5) [普通名詞] [カラ] [動詞] [レル/ラレル]
キー: 品詞 LIKE "動詞%"
AND
前方共起: (語彙素 = "から" AND 品詞 = "助詞-格助詞") ON 1 WORDS FROM キー
AND
前方共起: (品詞 LIKE "名詞-普通名詞%") ON 2 WORDS FROM キー
AND
後方共起: (語彙素 = "れる" OR 語彙素 = "られる") ON 1 WORDS FROM キー
WITH OPTIONS
unit="2" AND tglWords="50" AND tglKugiri="|" AND tglFixVariable="2"
(検索結果：13347 件)

(6) [普通名詞] [カラ] [(条件未指定語)] [ヲ] [動詞] [レル/ラレル]
キー: 品詞 LIKE "動詞%"
AND
前方共起: 語彙素 = "を" ON 1 WORDS FROM キー
AND
前方共起: (語彙素 = "から" AND 品詞 = "助詞-格助詞") ON 3 WORDS FROM キー
AND
前方共起: (品詞 LIKE "名詞-普通名詞%") ON 4 WORDS FROM キー
AND
後方共起: (語彙素 = "れる" OR 語彙素 = "られる") ON 1 WORDS FROM キー

WITH OPTIONS
unit="2" AND tglWords="50" AND tglKugiri="|" AND tglFixVariable="2"
（検索不能：検索開始後 20 分経過で中止）

検索件数から言って、カラ格名詞として普通名詞を指定した場合、人手による用例チェックは難しくなってくるだろう。用例をランダムに見て、目についた用例を拾うという使い方ならともかく、やはり構文的な情報を定量的に扱うのは、少なくとも本稿が試みた方法では難しいと言わざるを得ない。また現状の「中納言」の反応速度から言っても、検索開始から検索件数表示まで３～８分ほどかかっており、少なくともウェブのサービスとしてはこのくらいが限界とみたほうがよいだろう。

なお、参考までに、短単位検索の場合、上記４条件で抽出された件数は以下の通りである。

(7) ［固有名詞一般/人名］［カラ］［動詞］［レル/ラレル］
検索結果：269 件
(8) ［固有名詞一般/人名］［カラ］［(条件未指定語)］［ヲ］［動詞］［レル/ラレル］
検索結果：64 件
(9) ［普通名詞］［カラ］［動詞］［レル/ラレル］
検索結果：6952 件
(10) ［普通名詞］［カラ］［(条件未指定語)］［ヲ］［動詞］［レル/ラレル］
検索結果：575 件

いずれも長単位検索のほうがより多くの用例を抽出できており、ひとまずその意味に限っては、長単位検索のほうが構文論的な研究に向いているとは言えるかもしれない。

なお、上記［(条件未指定語)］について補足しておく。本来であれば、この部分、すなわち動詞からの距離が２語である部分は名詞を指定したいところではあるのだが、実際に検索条件として名詞を指定してしまうと、「中納言」から反応が返ってこなくなってしまう。これは、すべての名詞のチェックをしなければならなくなるために、検索コストが膨大なものになってしまっているということであろう。そのため、ここは条件としてなにも指定しないという消極的な語の指定方法によっている。それでも、後続の語の語彙素が「を」であることを指定してあれば、格助詞ヲが後接するのは名詞であるのはほぼ自明であるので、あえて検索条件として名詞で

あることを指定する必要はない、ということである。名詞の下位分類として固有名詞等を指定した場合の「中納言」の反応速度については未検証である。この場合、システム内部的に検査対象となる名詞の数も少なくなるはずなので、反応が返ってくるようになることも考えられる。これについては今後検証したい。

4. 分析

以下では、前節で検索された用例を用いて、若干の分析を行う。用例数が多く、紙幅の制限もあるので全般にわたった分析は今後の検討課題とせざるを得ないが、いくつかの点については注目すべき結果が得られている。

カラ格受動文に関しては、砂川 (1984) 以降、もっぱら「に」と「から」の使い分けという観点から分析が行われてきた。すなわち、ニ格受動文とカラ格受動文の違いはどのようなものか、あるいは、どのような場合にニ格受動文がカラ格受動文に言い換え可能か、といった視点が論考上の主要なポイントであった。そこには、受動文の動作主はニ格で表示されるのが基本であり、カラ格で表示されるのは特殊な場合である、ということが暗黙の了解事項として存在していたように思われる。

しかしながら、現実の用例を見る限り、カラ格受動文はかなり広範に用いられており、そのうちの一部には、後述するとおりニ格との交替が不可能なものも存在する。すなわち、ニ格かカラ格かという選択以前に、カラ格でなければならない用例が存在するのである。この「カラ格でなければならない受動文」についての検討は、先行研究の視点から抜け落ちていたように思われる。

以上のようなことから、ここでは、先行研究の検討事項を踏まえつつも、その延長線上に議論を展開することはいったん措いて、いま一度、検索された用例を慎重に検討する作業から始めたいと思う。

4.1. 検討対象となる用例について

用例の検討にあたっては、(3) および (4) の検索条件により抽出された用例を使用した[*2]。これらをすべて目視で確認し、カラ格受動文と見なし得ないものは分析の対象から除外した。この結果分析対象となったのは、(3) によるものが 633 例（全

*2 (5) (6) については、「中納言」自体の更新後に、あらためて調査・分析方法を検討したい。

検索用例の 87.8%)、(4) によるものが 136 例（同 80.0%）である。比較的高確率で目的とする用例が抽出できたと言ってよいだろう。除外した用例で多かったものは、カラ格名詞句が理由や純粋な起点を表すもの、レル/ラレルが尊敬あるいは自発をあらわすものであった。一部には解釈に迷ったものもあり、その場合も分析対象から外してある。分析の対象とした用例の動詞を本稿末に読みの五十音順で示す。括弧内の数字はその動詞が出現した用例数である。表記については、「中納言」が出力した語彙素の表記をそのまま記した[*3]。語彙素は表記揺れを吸収した概念なので、質的に同一の動詞が複数に分割されていることはない。

以下、これら動詞群のうちからいくつかの注目すべき類型を抽出し、その特徴について述べる。

4.2. 発行・発売等（発する動き）

以下の用例の動詞は、特定の場所において対象になんらかの位置づけを与えるといった意味を表す動詞である。この場合の動作主は状況の発生場所であるという意味で起点的であるとは言えるが、移動を表すわけではないという理由から着点の存在は含意されない。また、存在しなかった対象物が存在することになるという点では生産動詞的な側面を持っているとも言える。

(11) 刊行する、公表する、交付する、告示する、再刊する、再版する、策定する、出荷する、出版する、出版（覆刻）する、上梓する、創刊する、出す、発行する、発売する、発表する、発令する、復刊する、放映する、リリースする、等

以下これらの実例を示す。

(12) 同じく野上豊一郎の監修により昭和十七年から十九年にかけ創元社から刊行された『能楽全書』全六巻は、戦前の能楽研究の集大成とも言えるものだが、(田代慶一郎『謡曲を読む』朝日新聞社 1987)

(13) 新聞の発行部数は毎年，日本新聞協会から発表されており，また，普及率

*3 ただし、「きかされ」という表記の用例で、本来「聞かす」と解釈すべきところを「効かす」であると判断された例が２例あり、それらについては「聞かす」の用例数に含めた。

は新聞１部当りの人口あるいは１世帯当りの発行部数として計算されるものである。（佐藤研司/中山勝己『広告概論』ダイヤモンド社 1988）

これらのタイプは、先行研究においてはカラ格受動文の範疇には含まれていなかったように思われる。しかし、カラ格名詞句は動作主の意味役割を持っており、以下の通り対応する能動文においてはガ格で表示される。

(14) 創元社が『能楽全書』全六巻を刊行する
(15) 日本新聞協会が新聞の発行部数を発表する

このタイプの受動文においては、動作主をニ格で表示することは不可能である。

(16) *『能楽全書』全六巻が創元社に刊行される
(17) *新聞の発行部数が日本新聞協会に発表される

すなわち、このタイプの受動文では益岡(1987)の言うところの受影者を想定することができず、そのために動作主をニ格で表示することができないのだと考えられる。その点でこれらの受動文は受影性の関与しない降格受動文であると言える。

またこのタイプの受動文は、動作主をニヨッテで表示することが可能である。

(18) 『能楽全書』全六巻が創元社によって刊行される
(19) 新聞の発行部数が日本新聞協会によって発表される

この点も、降格受動文の特徴と一致する。

以上のように、このタイプのカラ格受動文は、ニ格かカラ格かという選択の結果動作主がカラ格によって表示されているわけではなく、そもそもカラ格によってしか動作主を表示できないのである。

4.3. 送付・授与等（相手先への位置の変更）

以下の用例の動詞は、動作主が相手に向けて対象物を動かす動きをあらわすもので、相手が着点としての意味役割を持ちうるものである。

(20) 送る、配る、加える、献上する、出願する、提出する、吐き出す、振り込

む、郵送する、授与する、譲渡する、送付する、提供する、派遣する、等

以下これらの実例を示す。

(21) 花鈴の生んだ女の子の身の上に起こったことを、旅行から帰った良は、留守番電話に入っていた母のメッセージと、花鈴から送られた何通かのｅメールで知った。(中上紀『パラダイス』恒文社 2001)

(22) 二十年中(二十年１月１日～９月三十日の間)に納付した額を証明する「社会保険料(国民年金保険料)控除証明書」が、社会保険庁から送付されます。(『市報とおかまち「だんだん」』2008 年 21 号 新潟県十日町市 2008)

(23) 新しい本が出版されて五日後、スタンペ男爵夫人から提供されたトーヴァルセンのアトリエのあるニーセェアにアンデルセンは行った。(アリソン・プリンス/黒田俊也『ハンス・クリスチャン・アンデルセン哀しき道化』愛育社 2005)

砂川(1984)も指摘しているとおり、この構文タイプにおける受動文には、相手を受動文主語としたものと、対象を受動文主語としたものの２パターンがある*4。

(24)

a. X[動作主]ガ Y[相手]ニ Z[対象]ヲ V〈移動〉

b. Y[相手]ガ X[動作主]カラ Z[対象]ヲ V〈移動〉レル

b.' Z[対象]ガ X[動作主]カラ Y[相手]ニ V〈移動〉レル

(21) (22) (23)であげた例はいずれも、(24) b.'の形、すなわち対象が受動文主語となったものである。この場合、ニ格とカラ格の交替可能性を検討してみると、以下の通り動作主をニ格で表示することは非常に難しい。

(25) ??花鈴に送られた何通かのｅメール(cf. ??何通かのｅメールが花鈴に送られた)

*4 ただし、砂川の動詞グループと本稿の動詞グループが同じものを指しているかどうかはさらに検討が必要である。

(26) ??「社会保険料（国民年金保険料）控除証明書」が、社会保険庁に送付されます

(27) ??スタンペ男爵夫人に提供されたトーヴァルセンのアトリエ（cf. ??トーヴァルセンのアトリエがスタンペ男爵夫人に提供された）

このことについて砂川 (1984) は、

この形は許容されないことが多い。「×田中さんに佐藤さんに紹介された」

と述べている。「紹介する」が果たして本稿が述べるこのタイプの動詞であるのかどうかについては慎重な判断が必要であろうが、少なくとも上記のような二重ニ格受動文が許容されにくいのはそのとおりであろう。この理由を本稿なりに解釈すれば、動作主がニ格の形式で表示されることによって、同じヒト性を持つ相手（それが名詞句として実現しているといないとにかかわらず）ニ格との弁別が難しくなっているために許容されにくい、というようなことであろうが、加えて本稿ではもうひとつ、受動文主語の意味素性についても触れておきたい。砂川 (1984) は受動文主語を明示していないため、少なくともこの例については明確なことは言えないが、(21) (22) (23) について言うなら、これらはいずれも無生物の対象が主語となったものであり、その点で、これらは降格受動文であって、そもそも受動文主語に対する受影性が働き得ず、したがって動作主をニ格の形式で表示する受影受動文としては実現できないのである。つまり、これらの受動文も、ニ格かカラ格かという選択の結果動作主がカラ格によって表示されているわけではなく、そもそもカラ格によってしか動作主を表示できないのである。

なお、さらに一点付け加えておくと、このタイプの受動文は、(21) (23) のように、実際には対象物を主名詞とする連体修飾構造として実現している場合が多い。このような場合、「受動」という概念は非常に希薄であり、むしろ属性叙述的であるとさえ言える。このことについては、今後より詳細な分析が必要である。

4.4. その他の動詞

その他の動詞では、動作主がニ格・カラ格のいずれでも問題ないようなものが多く認められた。この中には、従来からニ／カラ交替を起こすと言われてきた「言う」「教える」なども含まれている。以下、ニ格受動文・カラ格受動文のいずれもが可能な動詞のうち、比較的数多く検索されたものをあげる。

(28) 言う、依頼する、教える、聞かす、聞く、嫌う、誘う、叱る、知らす、信頼する、相談する、尋ねる、頼む、注意する、表彰する、褒める、任す、見せる、命ずる、求める、呼ぶ、等

以下これらの実例を示す。

(29) シカゴに戻ったら寄るようにと、ムーニーから言われていた。そこで、ムーニーの仲間“タフ・トニー”・カベツィオから借りた広いアパートに新妻を落ち着かせると、すぐオーク・パークへと車を走らせた。(チャック・ジアンカーナ/サム・ジアンカーナ/落合信彦『アメリカを葬った男』光文社 1992)

(30) その後すぐ、イプセンの芝居『ソルハウグの饗宴』につけた音楽をプフィッツナーから依頼され、その次はクリスマス・メルヘン『リューベツァール』を、たて続けに上演しなければならなかった。(ベルント・W・ヴェスリンク/香川檀『フルトヴェングラー』音楽之友社 1986)

(31) ウォーターゲート事件の時の録音テープの処分をニクソンから尋ねられて、「燃やしてしまえ」とニクソンに助言したのも彼だという。(副島隆彦『世界覇権国アメリカを動かす政治家と知識人たち』講談社 1999)

一方で、従来の先行研究からすれば非典型的とせざるを得ないようなタイプのカラ格受動文についても、いくつか興味深い用例が検索されている。

(32) あるいは、美少女高校生・麻生美樹の罠にはめられ、淫乱人妻コンビの麗華と真奈美から精液を搾り取られることに。もっとも、そんな事件に関わるうち、次第に恵と親密度を増していく。(堂本烈『ただいま淫交レッスン中！3(女子寮潜入編)』二見書房 2004)

(33) 六十六年から始まる後のプロレタリア文化大革命で劉少奇は、自らが用いたこの運動の手法によって、毛沢東から葬られることになる。(正木義也『台湾の悲劇』総合法令出版 2000)

(34) タクヤさんまで、ユミからツネられてたもの。最初は四人で泳いだり、遊んだりしてたんだが、そのうちに、自然とふた組のカップルになっちゃって。(『禁断の投稿ドキュメント』大洋図書 2004)

ここでは、4.2.および 4.3.で挙げた例において動作主マーカーとしてカラしか選択できない状況の中でカラ格が出現していたのと異なり、ニ・カラいずれもが選択可能な状況の中で、意識的にカラ格が選択されていると言ってよい。「使い分け」という表現で格形式の選択について述べるならば、それはここで取り上げたような用例について言うべきであろう。

これらの用例でカラ格が選択されている理由や、全体としてどのようなカラ格受動文のタイプが規定できるのかという点については、ここではこれ以上は立ち入らない。しかし、ひとつ注意すべきこととして、これらの用例において動作主マーカーとしてカラが選択されていることの理由として、先行研究で言われてきた「起点性を表示する」という理由だけでなく、「ニ格による受影性の強制を避ける」という理由があるのではないかということを指摘しておきたい。

例えば、「言う」「教える」などの受動文においては、相手が単純に情報を受け止めたという以上に、そのような動作を受けることによって、結果的に相手がなんらかの情報を保持した状態に変化するという解釈が可能であり、その点で、相手に対する影響性が明確である。そのような場合には、受影受動文として動作主をニ格の形式で表示することが可能である。その一方、ニ格受動文に伴うそうした受影性の解釈の強制が、表現の意図の上で、発話者ないし書き手にとって余分なものとなってしまうこともあり得るだろう。語用論的な影響性の解釈を省き、単純な事態として受け手側から事象を記述したいこともあるはずである。そのような場合に、受影性の解釈をキャンセルする方法として、カラ格を選択するという用法が広がりを見せつつあるのではないかと思われるのである。当然、格形式の選択には複合的な要因が関係しており、すべてをひとつのルールで説明することは不可能であろう。しかしながら、受影性のキャンセルという観点は、カラ格受動文の用例を実際に見ていくと、十分検討するに値するように思われる。この点を含め、カラ格受動文全体のタイプ分けとその生起の原則についての検討は、今後の課題としたい。

以上、ここでは、カラ格受動文からニ格受動文への変換可能性について見ることにより、カラ格形式が選択されている理由の考察を試みた。要点は以下の 2 点である。

(35) 降格受動文であり、そもそもニ格形式の選択の余地がないタイプのカラ格受動文が存在する

(36) ニ格・カラ格両方の選択の余地がある場合、受影性の解釈をキャンセルするためにカラ格を選択する可能性がある

5. おわりに

本稿では、前半で「中納言」を用いた用例検索の具体的な手法について述べ、後半で、実際に検索されたカラ格受動文の用例について若干の分析を試みた。今回用いた用例は、カラ格名詞句と動詞が直接接続しているもの、およびカラ格名詞句と動詞の間にヲ格名詞句が挟まっているものだけであり、すでに述べたとおりこれがすべての用例ではないので、より多くのパターンのカラ格受動文を抽出する技術的な方法を検討しつつ、カラ格受動文全体についての分析を進めていきたい。

筆者は、和氣(2006)においても、〈資格〉の意味を表すニ句[*5]を抽出するために、『CD-ROM 版新潮文庫の 100 冊』テキストデータからの用例検索を行った。その際に用いた方法は、以下のような正規表現を用いてテキストを grep するというようなものである。

(37) [亜-腕]+を[亜-腕]+に[亜-腕]+[あ-ん]+

これはすなわち、JIS 第一水準の漢字がひとつ以上連続したものを名詞および動詞語幹の可能性があるものと見なし、そこに格助詞ヲとニが挟まり、最後に送り仮名が付されているようなパターンの抽出を狙ったものであった。これで検索されたのが 3275 例、うち、目視により資格のニ句が現れている構文と判断したのが 107 例（約 3.27%）であった。異なる構文であるので単純な比較はできないにせよ、「中納言」を利用することによって品詞や一部の名詞の意味まで指定しつつ複雑な構文パターンを抽出できることが、用例検索においていかに有効であるかは理解できるだろう。計算機を利用して構文的あるいは意味的な条件をもとに用例を自由自在に検索することは、現在の自然言語処理技術ではまだ難しいといわざるを得ないが、現在実現されている技術を利用することで、限定的であるにせよ、用例検索が一定程度有効に行えることは示し得たと考える。

---

*5 〈資格〉のニ句とは、以下の例の「支配人に」のような、結果的にヲ格名詞（以下では「内務省衛生局の役人であった者」）に対して与えられる位置づけを表すものである。

○内務省衛生局の役人であった者を支配人に迎え、その官庁内への顔を最大限に利用し、各種の便宜を得たりしていた。（星新一『人民は弱し　官吏は強し』）

参考文献

杉本武 (2007)『現代日本語の受動文と格の研究』筑波大学学位請求論文 博士（言語学） 筑波大学

砂川有里子 (1984)「「ニ」と「カラ」の使い分けと動詞の意味構造について」『日本語・日本文化』12 pp.71-87 大阪外国語大学研究留学生別科

砂川有里子・井上優・小林ミナ・滝沢直宏・投野由紀夫・山内博之 (2010)『代表性を有する書き言葉コーパスを利用した日本語教育研究』文部科学省科学研究費補助金特定領域研究「日本語コーパス」日本語教育班 平成 20-21 年度研究成果報告書

田野村忠温・服部匡・杉本武・石井正彦 (2010)『コーパス日本語学の新展開』文部科学省科学研究費補助金特定領域研究「日本語コーパス」日本語学班 研究成果報告書

益岡隆志 (1987)『命題の文法』くろしお出版

森雄一 (1997)「受動文の動作主マーカーとして用いられるカラについて」『茨城大学人文学部紀要人文学科論集』30 pp.83-99 茨城大学

和氣愛仁 (1997)「文法的ヴォイスとニ格名詞句」『筑波日本語研究』2 pp.57-73 筑波大学文芸・言語研究科 日本語学研究室

和氣愛仁 (2006)「〈資格〉のニ句について」矢澤真人・橋本修編『現代日本語文法現象と理論のインタラクション』pp.169-190 ひつじ書房

Uchimoto, K., & Isahara, H. (2007). Morphological annotation of a large spontaneous speech corpus in Japanese, Proceedings of IJCAI, pp.1731-1737

ウェブ上のリソース

「中納言」
https://chunagon.ninjal.ac.jp/

「少納言」
http://www.kotonoha.gr.jp/shonagon/

科学研究費特定領域研究「代表性を有する大規模日本語書き言葉コーパスの構築」
http://www.tokuteicorpus.jp/

国立国語研究所言語コーパス整備計画 KOTONOHA
http://www.ninjal.ac.jp/kotonoha/

UniDic

http://www.tokuteicorpus.jp/dist/

naist-jdic

http://sourceforge.jp/projects/naist-jdic/

ChaSen

http://chasen-legacy.sourceforge.jp/

MeCab

http://mecab.sourceforge.net/

CaboCha

http://code.google.com/p/cabocha/

ChaKi

http://sourceforge.jp/projects/chaki/

ひまわり

http://www2.ninjal.ac.jp/lrc/

検討対象とした動詞

a. [固有名詞一般/人名] [カラ] [動詞] [レル/ラレル]

愛す (1)、握手する (1)、預ける (1)、与える (12)、言い含める (2)、言う (18)、諫める (1)、委嘱する (3)、委託する (3)、依頼する (8)、受け入れる (1)、疑う (1)、打ち明ける (1)、訴える (2)、生み出す (1)、追う (2)、送り込む (1)、送る (27)、贈る (20)、行う (1)、怒る (3)、教える (16)、押し付ける (1)、押す (2)、御願いする (1)、回付する (1)、介抱する (1)、掛ける (3)、貸し与える (1)、下賜する (1)、噛み付く (1)、からかう (1)、可愛がる (1)、監禁する (1)、刊行する (20)、勘当する (1)、勧誘する (1)、聞かす (31)、聞く (5)、寄進する (2)、寄贈する (2)、期待する (1)、義務付ける (1)、糾問する (1)、拒絶する (2)、拒否する (1)、嫌う (3)、切り売りする (1)、下す (1)、配る (2)、加える (1)、啓発する (1)、軽蔑する (1)、激賞する (1)、激励する (1)、蹴る (1)、献上する (2)、請い受ける (1)、請う (1)、厚遇する (1)、攻撃する (1)、公表する (1)、交付する (3)、告示する (1)、鼓吹する (1)、再刊する (1)、催促する (1)、再版する (1)、削除する (1)、策定する (1)、差し入れる (1)、授ける (4)、誘う (4)、叱る (4)、支給する (2)、指示する (1)、叱責する (1)、質問する (1)、指定する (2)、指摘する (2)、支払う (1)、指名する (1)、示す (2)、出荷する (1)、出願する (1)、出

題する(1)、出版する(33)、出版(覆刻)する(1)、授与する(2)、背負い込まする(1)、紹介する(11)、上梓する(1)、招待する(1)、譲渡する(1)、招聘する(1)、奨励する(1)、除名する(1)、知らす(7)、尋問する(1)、信用する(1)、信頼する(3)、スカウトする(1)、勧める(2)、請求する(2)、制限する(1)、せがむ(1)、折檻する(1)、絶交する(1)、絶賛する(1)、折衷主義者呼ばわりする(1)、説得する(1)、説明する(1)、宣告する(1)、創刊する(2)、相談する(3)、送付する(5)、注ぐ(1)、託す(2)、出す(16)、尋ねる(4)、頼む(22)、注意する(4)、忠告する(1)、注文する(1)、通知する(1)、使わす(1)、告げる(3)、伝える(2)、突っ込む(1)、抓る(1)、提案する(1)、提供する(6)、提示する(4)、提出する(5)、手解きする(1)、手渡す(3)、伝授する(3)、討伐する(1)、咎める(1)、特命する(1)、届ける(3)、怒鳴り付ける(1)、怒鳴る(1)、飛ぶ(1)、流す(1)、慰める(1)、投げ掛ける(1)、投げ付ける(1)、投げる(2)、憎む(2)、睨む(2)、認可する(2)、任命する(1)、狙う(2)、望む(1)、配布する(1)、吐き出す(1)、派遣する(12)、発給する(1)、発行する(9)、罰する(1)、抜擢する(1)、発売する(22)、発表する(16)、発令する(1)、話し掛ける(1)、撥ね付ける(1)、判断する(1)、反論する(1)、引っ張る(1)、否定する(1)、非難する(2)、批判する(1)、表彰する(4)、復刊する(1)、振り込む(2)、プレゼントする(3)、プロポーズする(1)、返還する(1)、弁護する(1)、放映する(1)、報告する(4)、放出する(1)、放送する(1)、奉納する(1)、葬る(1)、褒める(3)、マークする(1)、任す(3)、招く(2)、見捨てる(1)、見せる(4)、認める(1)、見放す(1)、向ける(1)、命ず(2)、命ずる(4)、命令する(2)、召し寄せる(1)、申し入れる(1)、申し渡す(1)、齎す(1)、持て囃す(1)、遣り込める(1)、言う(2)、郵送する(1)、譲る(1)、許す(2)、要求する(2)、要請する(1)、呼び出す(5)、呼び止める(1)、呼ぶ(5)、離縁する(1)、リリースする(1)、渡す(18)

b. [固有名詞一般/人名][カラ][(条件未指定語)][ヲ][動詞][レル/ラレル]
与える(2)、あてがう(1)、安堵する(1)、言い渡す(2)、言う(3)、委譲する(1)、委託する(1)、挑む(1)、委任する(2)、依頼する(6)、打ち明ける(1)、打つ(1)、移す(1)、追い出す(1)、贈る(1)、押す(1)、落とす(1)、負わす(1)、返す(1)、掛ける(9)、下賜する(1)、課す(1)、聞かす(2)、聞く(1)、拒絶する(1)、禁ずる(1)、下す(1)、献上する(1)、断る(2)、差し向ける(1)、指示する(1)、指導する(1)、絞り取る(1)、示す(1)、授与する(4)、紹介する(3)、賞する(1)、勧める(1)、請求する(1)、制限する(1)、託す(3)、尋ねる(1)、叩き込む(1)、叩く(1)、立てる(1)、注入する(1)、付ける(2)、告げる(1)、停止する(2)、手渡す(4)、説く(1)、取り消す(1)、取る(1)、

投げ付ける(1)、労う(1)、発する(1)、バトンタッチする(1)、引き合わす(1)、引っ張る(1)、振り掛ける(1)、プレゼントする(1)、仄めかす(1)、任す(1)、守らする(1)、見せる(2)、命ずる(7)、申し入れる(1)、持たす(1)、持ち掛ける(1)、求める(36)、譲る(4)、許す(2)、要請する(1)、渡す(7)

わき　としひと／人文社会系
(2011年11月30日 受理)