

# 通時コーパス用『中納言』： Web ベースの古典語コンコーダンサー

小木曾 智信 (国立国語研究所言語資源研究系)<sup>†</sup>  
中村 壮範 (マンパワージャパン株式会社)

## *Chunagon* for the NINJAL Diachronic Corpus: a Web-based Concordancer of Classical Japanese

Toshinobu Ogiso (National Institute for Japanese Language and Linguistics)  
Takenori Nakamura (Manpower Japan Co., Ltd.)

### 1. はじめに

国立国語研究所の共同研究プロジェクト「通時コーパスの設計」<sup>1</sup>では、日本語の歴史的資料をコーパス化するための研究が行われている。その一環として、先行して整備が進んでいる一部のデータを格納した Web アプリケーション・通時コーパス用『中納言』の共同研究者向けの公開を開始した。通時コーパス用『中納言』は『現代日本語書き言葉均衡コーパス』(以下 BCCWJ とする)の公開にあたって開発された Web ベースのコンコーダンサー『中納言』に若干の機能拡張を行い、『源氏物語』などの通時コーパスの一部のデータを格納したものである。これにより、一般の古典研究者にも使いやすいインターフェイスを用いて通時コーパスを利用することが可能になった。本発表では通時コーパス用『中納言』と現在利用可能なデータについて紹介する。

### 2. 通時コーパス用『中納言』の概要

『中納言』はコーパスに付与された形態論情報を用いて高度な検索を行うことが可能な Web アプリケーションである。検索条件指定の自由度が高く、複数の語を組み合わせ、詳細な条件指定を行うことができる。検索結果はキーとなる語の形態論情報、サンプルの書誌情報とともに KWIC 形式で一覧表示されるほか、表形式のテキストデータとしてダウンロードして利用することもできる。

今回、準備中の通時コーパスデータをこのシステムに格納して利用することを可能にした(次ページ図1)。BCCWJでは「短単位」と「長単位」の二つの異なるサイズの形態論情報を用いることができたが、通時コーパスのデータでは今のところ長単位の整備が進んでいないため、通時コーパス用『中納言』では短単位だけが利用可能となっている。

通時コーパス用『中納言』では、上記『中納言』の機能に加えて BCCWJ にはなかったテキストに関する情報が利用できるようになっている。一つは「本文種別」と呼んでいる情報で、小学館新編古典文学全集に付けられている情報を元に、「会話」「手紙」「歌」「詞書」とそれ以外(地の文)の別が、個々の語について付与されている。さらに「話者情報」として、「会話」についてはその話者、「手紙」については書き手、「歌」については歌番号などの情報が付与されている。

---

<sup>†</sup> togiso@ninjal.ac.jp

<sup>1</sup> <http://historicalcorpus.jp>



図1 通時コーパス用「中納言」検索実行画面

### 3. 収録データ

現在、通時コーパス用『中納言』に格納され利用可能になっているデータは表1に示した13作品、約87万語である。これらの作品は、後述する「中古和文 UniDic」による自動形態素解析結果をもとに、すべて一度は人手による修正・チェックを経たものである（ただし、一部データの抜き取り調査によると現時点での精度はおおむね98%程度であると思われる）。国語研究所の通時コーパスは現在のところ設計の途上であり、ごく一部のデータが作成されているに過ぎない。それでも、『源氏物語』をはじめとする中古の主要な古典文学作品をカバーしている。

このうち『竹取物語』『伊勢物語』『土佐日記』『大和物語』『枕草子』『源氏物語』の6作品は、小学館の新編日本古典文学全集に基づくデータであり、最終的な通時コーパスにも同じテキストが用いられる予定である。残る7作品は、入手しやすいデータをもとに形態素解析の学習用データとしたものや、研究・試験用に作成したものであり、参考データにとどまる。

表 1 収録データ (2012年6月30日現在)

作品名	語数 (短単位)	備考
竹取物語	12583	小学館・新編日本古典文学全集
伊勢物語	15900	
土佐日記	8113	
大和物語	26733	
枕草子	79879	
源氏物語	510714	
古今仮名序	3107	その他 (「中古和文 UniDic」学習用データ)
紫式部日記	20346	
大鏡	82796	
更級日記	16652	
方丈記	4191	
徒然草	41675	
恋路ゆかしき大将	44819	
計	867508	

#### 4. 形態論情報

通時コーパスのデータは、BCCWJと同様に、形態素解析技術を用いて全ての本文テキストに単語の切れ目、読み、品詞、活用などの形態論情報を付与している。形態素解析のための辞書は、BCCWJの構築に用いられた「UniDic」をもとに、中古和文を解析できるように語彙を増補しパラメータを調整した「中古和文 UniDic」を用いている。「中古和文 UniDic」は、未知語のないテキストであればおおむね96～97%程度の精度で解析を行うことが可能になっている<sup>2</sup>。

「中古和文 UniDic」が付与する形態論情報は、BCCWJと同様の「短単位」を採用し、中古語であっても現代語とできるかぎり基準を揃え、相互に比較することができるように配慮したものである。ただし、語の歴史的变化や中古語の実態を踏まえ、時代別に異なった扱いをしている語も少なくない。たとえば、現代語では連体詞とされる「この」「その」が、中古語では代名詞「こ」「そ」と格助詞「の」に分けて数えられている。この中古和文用の短単位の規定は、小椋・須永(2012)にまとめられている<sup>3</sup>。

通時コーパス用『中納言』を用いて中古語の検索をする場合には、この短単位の規定について理解をしておく必要がある。

#### 5. 検索方法

『中納言』に格納されているデータは形態論情報が付与されているため、表層の文字列だけでなく、形態論情報を利用することで高度な検索条件の指定を行うことができる。たとえば、語彙素「給う」(終止形)を指定することで「給う」「給は」「給ひ」「給ふ」などの各活用形を一括で検索することが可能である。また、UniDicの見出し語の階層構造により、見出し語を語彙素で指定すれば、その異表記を一括検索することができる。したがって、漢字表記と仮名表記の違い、異体字や送り仮名の揺れなどを一々意識することなく検索できる。

<sup>2</sup> 「中古和文 UniDic」は次のサイトで報告書 PDF と共に一般公開している (無償)。

<http://www2.ninjal.ac.jp/lrc/index.php?UniDic>

<sup>3</sup> 『中古和文 UniDic 短単位規程集』の PDF ファイルも上記ウェブサイトで一般公開している。

## 5. 1 検索条件の指定

具体的には、形態論情報を使った検索では、次のコントロールで検索条件を設定する。

キー ( --- 10 語 )  
---選択--- が [ ] [ ] 短単位の条件の追加

「選択」で条件指定する属性（「語彙素」「出現書字形」など）を選び、右の空欄でその中身を指定する。

キー ( --- 10 語 )  
語彙素 が 読む [ ] [ ] 短単位の条件の追加

「短単位の条件の追加」ボタンで一つの単位について詳細な条件指定を追加できる。次の例では、語彙素が「読む」でかつ、活用形が「連体形」の例を検索している。（活用形など選択肢が決まっているものはドロップダウンメニューから選択する）

キー ( --- 10 語 )  
語彙素 が 読む [ ] [ ] 短単位の条件の追加  
AND 活用形 の 大分類 が 連体形 [ ] [ ] 短単位の条件の追加

さらに、複数の単位を組み合わせることもできる。「前方共起条件の追加」ボタンでキーの前方に出現する単位を指定、「後方共起条件の追加」ボタンでキーの後方に出現する単位を指定する。共起条件は前方後方合わせて最大10個まで追加できる。

共起位置は、「キーから」または「文頭から」を基準として、「n語」または「n語以内」のように、細かく指定することができる。次の例は、「美しい」の連体形の直後（後方1語）に来る名詞を検索したものである。

前方共起1 ( キーから 1 語 )  
語彙素 が 美しい [ ] [ ] 短単位の条件の追加  
AND 活用形 の 大分類 が 連体形 [ ] [ ] 短単位の条件の追加  
キー ( --- 10 語 )  
品詞 の 大分類 が 名詞 [ ] [ ] 短単位の条件の追加  
▲ 後方共起条件の追加

『中納言』では、形態論情報を使った検索以外に「文字列検索」によって表層の文字列にもとづく検索を行うこともできる。この場合にも、検索結果は形態論情報付きで表示されるため、調査したい語にどのような形態論情報が付与されているか分からない場合には、いったん文字列検索で形態論情報を確認すると便利である。

なお、このようにして画面上で指定した検索条件は、システムが解釈できる「検索条件式」に変換されたのち、検索が実行される。この検索条件式は、検索履歴として自動的にサーバー上に記録されるほか、画面上で編集をして再検索に利用することが可能になっている。

たとえば上記の前方共起1を利用した検索例は、次の検索条件式で表される。

```
キー: 品詞 LIKE "名詞%" AND 前方共起: (語彙素 = "美しい" AND 活用形 LIKE "連体形%")  
ON 1 WORDS FROM キー IN core="true" OR core="false" WITH OPTIONS unit="1" AND  
tglWords="20" AND tglKugiri="|" AND tglFixVariable="2"
```

この条件式を控えておくことにより、『中納言』のユーザーであれば全く同じ検索を再現することができる。

## 5. 2 検索実行とダウンロード

検索の実行には画面上の「検索」ボタンをクリックする。これにより、画面下部に図 2 のような検索結果が表示される。ただし、画面上に表示される用例数は 500 例までとなっている。これ以上の用例を確認する場合には結果をダウンロードする必要がある。

24 件の結果が見つかりました。

☐ テーブルの幅を固定 短

サンプル ID	前文脈	キ	後文脈	語彙素読み	語彙素	語彙素細分類	語形	品詞	活用型	活用形	サブコーパス名	執筆者	書名 / 出典	出版年	本文種別	話者	文件
20_源氏物語 02_19 薄雲	心やすく産かしき心ばまぬれば、上 にいとよくつき寝ひきこえたまへれば、いみじう(うつしき)	もの	御たりと思しけり。他ごとなく抱き抱ひ、もてあそび居こえたまひて、乳母もあつからば(近う)	モノ	物		モノ	名詞-普通名詞-サ変可能			中古和文コア	紫式部	源氏物語	1010			
22_源氏物語 04_35 若菜下	ところもなくおまゆも、はすがりに、腹あしくてもねたみぢちたる、腰敬げきて(うつしき)	人	さまにこそものしたまふ(ゆる)。 院は、 対へ 渡りたまひぬ。上 は、	ヒト	人		ヒト	名詞-普通名詞-一般			中古和文コア	紫式部	源氏物語	1010			
24_源氏物語 05_49 宿木	ほどはあるにまかせてはいらかならんと思ひまてて、いとわづらひに、(うつしき)	さま	にもてなして(あたまへ)れば、いとど(あはれ)は、 わしく思され、 日ごろの(意り)	サマ	様		サマ	名詞-普通名詞-一般			中古和文コア	紫式部	源氏物語	1010			
22_源氏物語 04_37 横笛	と、いはいは御気色ゆかし。宮たちには、思ひなしこそ(気高)けれ、世の(常)の(うつしき)	児	どもと見えたまふ(こ)に、この(君)は、いとあてなるもの(から)、はま こと(こ)	チゴ	稚児		チゴ	名詞-普通名詞-一般			中古和文コア	紫式部	源氏物語	1010			
22_源氏物語 04_36 柏木	薫がちなる(今)様色など(着)たまひて、まだありつかぬ(御)かたは(は)ら目、 かく(て)しも(うつしき)	子	どもの心地して、ぬまめか(う)をかし(が)なり。 「い、 あな(心)憂、 墨染に(そ)、	コ	子		コ	名詞-普通名詞-一般			中古和文コア	紫式部	源氏物語	1010			

図 2 検索結果の表示

この検索結果の各列は、【列の表示】のチェックボックスにより表示の ON/OFF を切り替えることができる。検索結果のサンプル ID をクリックすることにより、当該位置周辺(前後 30 単位)に付けられている形態論情報を確認することができる(図 3)。

サンプル ID	前文脈	キ	後文脈	補助記号-句点	記号	中古和文	執筆	書名	出版	本文	話者	文件	
18_枕草子 138570	。		。	補助記号-句点	記号。	中古和文	1	1	22861.9				
18_枕草子 138580	いみじう	イミジイ	いみじい	形容詞-一般	文語形容詞-シク	連用形-ウ音便	イミジウ	和	いみじう	中古和文	1	1	22865.6
18_枕草子 138590	うつしき	ウツクシイ	美しい	形容詞-一般	文語形容詞-シク	連体形-一般	ウツクシキ	和	うつしき	中古和文	1	1	22870.5
18_枕草子 138600	ちご	チゴ	稚児	名詞-普通名詞-一般			チゴ	和	ちご	中古和文	1	1	22872.8
18_枕草子 138610	の	ノ	の	助詞-格助詞			ノ	和	の	中古和文	1	1	22873.9
18_枕草子 138620	いちご	イチゴ	苺	名詞-普通名詞-一般			イチゴ	和	いちご	中古和文	1	1	22876.7
18_枕草子 138630	など	ナド	など	助詞-副助詞			ナド	和	など	中古和文	1	1	22878.8
18_枕草子 138640	食ひ	クウ	食う	動詞-一般	文語四段-ハ行	連用形-一般	クイ	和	食ひ	中古和文	1	1	22880.8
18_枕草子 138650	たる	タル	たり	完了	助動詞	文語助動詞-タリ-完了	タル	和	たる	中古和文	1	1	22882.8
18_枕草子 138660	。		。	補助記号-句点	記号。	中古和文	1	1	22883.9				

図 3 キー周辺の形態論情報の表示

「検索」ボタンの代わりに「検索結果をダウンロード」ボタンをクリックすることによって、検索結果と検索条件式をテキストデータとしてダウンロードすることができる。データは zip 圧縮されており、アーカイブ中のファイル kwic.txt が検索結果のデータ(タブ区切りの表形式テキスト、文字コードは UTF-8)、summary.txt が検索条件式のデータとなっている。

## 6. 検索例

検索例として、完了の助動詞「つ」「ぬ」に上接する動詞のリストを検索する例を示す(図4)。ここでは、助動詞の前方2語以内に来る動詞をキーとして検索結果を取得している。

図4 助動詞「つ」の上接動詞の検索条件指定

この条件での検索結果の表示は図5のようになる。

2485 件の結果が見つかりました。そのうち 500 件を表示しています。 ☐ テーブルの幅を固定 短

サンプルID	前文脈	キー	後文脈	語彙素読み	語彙素分類	語形	品詞	活用型	活用形	サブコーパス名	執筆者	書名/出典	出版年	本文種別	話者	文体
1203_大和物語	て山にに入りけり。はてはこせたりける。いからくして思ひわする忍しさを捨て	鳴き	(つる)ひぐひすの声(こし、はては)君けすれ(かり)ひぐひすの(鳴く)をりのみ(や)思ひ(づ)べき	ナク	鳴く	ナク	動詞-一般	文語四段-力行	連用形-一般	中古和文コア		大和物語	951	歌		
1101_古今和歌集	とぶらまほむち葉の散りて(つ)もれる(わ)か(や)ど(に)たれ(を)まつ虫(に)こ(ろ)鳴(く)ら(む)び(ぐ)らし(の)鳴(く)	鳴き	(つる)なへ(こ)日(は)薄れぬ(と)思へ(ば)山(の)障(こ)を(あ)り(け)る(む)らし(の)鳴(く)	ナク	鳴く	ナク	動詞-一般	文語四段-力行	連用形-一般	中古和文コア	(一七七八)	古今和歌集	906	歌	歌番号(204)	
22_源氏物語 04_35若菜下	して、言はれ(た)方(な)く(ぬ)ま(ま)の(年)ご(う)、ま(め)事(に)も(あ)だ(事)に(も)召(し)ま(つ)ま(し)、(参)り	馴れ	(つる)もの(を)に、(人)は(り)ま(に)ま(や)か(に)思(し)と(ど)め(た)る(御)気(色)の(あ)ま(れ)に(な)つ(か)し(き)を、(あ)さ(ま)し(く)	ナレ	慣れる	ナル	動詞-非自立可能	文語下二段-う行	連用形-一般	中古和文コア	紫式部	源氏物語	1010			
23_源氏物語 06_52蜻蛉	も(あ)ら(ぬ)ば、(身)を(あ)げ(た)ま(へ)ら(ん)と(も)思(ひ)も(あ)ら(ず)、(鬼)や	食ひ	(つ)ら(ん)、(狐)め(く)も(の)や(と)り(も)て(去)ぬ(ら)ん、(し)と(昔)物(語)の(あ)や(し)き(も)の(事)	クウ	食う	クウ	動詞-一般	文語四段-ハ行	連用形-一般	中古和文コア	紫式部	源氏物語	1010			
20_源氏物語 02_13明石	い(か)で(思)ふ(は)ま(こ)に(見)た(て)ま(つ)ら(む)と、(年)月(を)頼(み)通(く)し、(今)や(思)ひ(か)な(ふ)に(と)そ(こ)	頼み	(つ)れ、(ま)め(人)の(心)変(る)ま(な)ご(り)な(ぬ)む(と)聞(き)し(し)ま(ま)こ(と)な(り)け(り)、(と)世(を)	タノム	頼む	タノム	動詞-一般	文語四段-マ行	連用形-一般	中古和文コア	紫式部	源氏物語	1010	会話		
22_源氏物語 04_39夕霧	つ(り)た(ま)ふ(れ)ど、(三)条(殿)、限(り)な(め)り(と)、(は)し(も)や(り)ま(に)そ(か)つ(は)し	頼み	(つ)れ、(ま)め(人)の(心)変(る)ま(な)ご(り)な(ぬ)む(と)聞(き)し(し)ま(ま)こ(と)な(り)け(り)、(と)世(を)	タノム	頼む	タノム	動詞-一般	文語四段-マ行	連用形-一般	中古和文コア	紫式部	源氏物語	1010			

図5 助動詞「つ」の上接動詞の検索結果

上記の画面は「つ」の上接動詞を検索するものだが、検索条件式中で括弧と OR 演算子を用いることで二つの助動詞を一度に検索するようにすることもできる。

キー: 品詞 LIKE "動詞%" AND 後方共起: ((語彙素 = "つ" OR 語彙素 = "ぬ") AND 品詞 LIKE "助動詞%") WITHIN 2 WORDS FROM キー IN core="true" OR core="false" WITH OPTIONS unit="1" AND tglWords="20" AND tglKugiri="|" AND tglFixVariable="2"

### 6. 1 検索結果の集計

このような検索によって取得したデータをダウンロードすることで、アプリケーションに読み込んでさまざまな処理を施し、集計やデータの分析に利用することができる。たとえば、Microsoft Excel のピボットテーブル機能を利用することで、高度な集計処理を容易に実現することができる。

図6は、上記の条件で検索した助動詞「つ」と「ぬ」の上接動詞のリストを Excel に読み込んで助動詞の情報を加えたものである。

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	助動詞	サンプルID	連番	前文脈	キ	後文脈	語彙素統	語彙素	語彙素細	語形	品詞	活用型	活用形
2	つ	24_源氏物語05_49宿木	31810	人を見なく見なしきこえ	たまう	てしに思ふにには、 箭の	タマウ	給う	尊敬	タマウ	動詞-非文語四段-連用形	ウ	ウ
3	つ	20_源氏物語01_04夕顔	94060	時たち離れたてまつらず	馴れ	きこえ(つる)人 に はか にナレル	慣れる			ナル	動詞-非文語下二形-連用形	ニ	ニ
4	つ	20_源氏物語01_04夕顔	94070	ち離れたてまつらず	馴れ	きこえ(つる)人 に はか にナレル	慣れる			ナル	動詞-非文語下二形-連用形	ニ	ニ
5	つ	20_源氏物語02_14薄櫻	4100	ためにには、いま見出で	たまひ	てしに思ふも 口 置しや	タマフ	給う	尊敬	タマフ	動詞-非文語四段-連用形	ニ	ニ
6	つ	20_源氏物語02_14薄櫻	7840	阿闍梨に よりて 位 を 返し	たてまつり	てしに思ふも 口 置しや	タマフ	給う		タマツル	動詞-非文語四段-連用形	ニ	ニ
7	つ	20_源氏物語01_04夕顔	75020	阿闍梨の せよ と 言ひ	やり	てしに思ふも 口 置しや	タマツル	奉る		ヤル	動詞-非文語四段-連用形	ニ	ニ
8	つ	20_源氏物語02_14薄櫻	22570	出で立る て、 學 の 心 地	見	てしに思ふも 口 置しや	ミル	見る		ス	動詞-非文語上二形-連用形	ニ	ニ
9	つ	20_源氏物語02_14薄櫻	39170	五筋を 思 し 忘れ ず、 また	見	てしに思ふも 口 置しや	ミル	見る		ス	動詞-非文語上二形-連用形	ニ	ニ
10	つ	20_源氏物語02_14薄櫻	53640	近き 惟光 に けたま は り や	し	てしに思ふも 口 置しや	ミル	見る		ス	動詞-非文語上二形-連用形	ニ	ニ
11	つ	20_源氏物語01_04夕顔	75410	もの せ させたまふ に と や	はべり	てしに思ふも 口 置しや	ミル	見る		ハベリ	動詞-非文語上二形-連用形	ニ	ニ
12	つ	20_源氏物語02_14薄櫻	60460	聞こえあはせ 人 に 思ひ	きこえ	てしに思ふも 口 置しや	キコエル	聞こえる		キコユ	動詞-非文語下二形-連用形	ニ	ニ
13	つ	20_源氏物語02_14薄櫻	62060	たてまつら と に こ 思ひ	たまへ	てしに思ふも 口 置しや	タマフ	給う	謙讓	タマフ	動詞-非文語下二形-連用形	ニ	ニ
14	つ	20_源氏物語02_14薄櫻	74610	か で さ やか に 御容 を	見	てしに思ふも 口 置しや	ミル	見る		ミル	動詞-非文語上二形-連用形	ニ	ニ
15	つ	20_源氏物語02_14薄櫻	82360	ゆ やう に て こ に 還し	たてまつり	てしに思ふも 口 置しや	タマツル	奉る		タマツル	動詞-非文語四段-連用形	ニ	ニ
16	つ	21_源氏物語03_24胡蝶	20000	文 大臣 に も 知 せ し	し	てしに思ふも 口 置しや	ス	知る		ス	動詞-非文語上二形-連用形	ニ	ニ
17	つ	21_源氏物語03_24胡蝶	35950	もの ほど も、 見 あ は し	はて	たまひ(て)も や など、	果てる			ハツ	動詞-非文語下二形-連用形	ニ	ニ
18	つ	21_源氏物語03_24胡蝶	35960	もの ほど も、 見 あ は し	たまひ	たまひ(て)も や など、	果てる		尊敬	タマフ	動詞-非文語四段-連用形	ニ	ニ
19	つ	21_源氏物語03_24胡蝶	39350	思し より て、 口 の の 心	得	てしに思ふも 口 置しや	ミル	得る		ウ	動詞-非文語下二形-連用形	ニ	ニ
20	つ	20_源氏物語01_04夕顔	61530	弱 く て、 思 も 空 を の み	見	てしに思ふも 口 置しや	ミル	見る		ウ	動詞-非文語上二形-連用形	ニ	ニ
21	つ	21_源氏物語03_24胡蝶	43350	た き に、 思 も こ と な く て	過ぎ	てしに思ふも 口 置しや	スグ	過ぎる		スグ	動詞-非文語下二形-連用形	ニ	ニ
22	つ	22_源氏物語04_34若菜上	530	本意 深き を、 后 の 宮 の	おほしまし	てしに思ふも 口 置しや	オワシマス	おわします		オワシマス	動詞-非文語四段-連用形	ニ	ニ
23	つ	22_源氏物語04_34若菜上	8590	な や ませたまふ に と も	あり	てしに思ふも 口 置しや	アル	有る		アリ	動詞-非文語上二形-連用形	ニ	ニ
24	つ	22_源氏物語04_34若菜上	8940	の し たり、 御 位 を 去 らせ	たまひ	てしに思ふも 口 置しや	タマフ	給う	尊敬	タマフ	動詞-非文語四段-連用形	ニ	ニ
25	つ	22_源氏物語04_34若菜上	12280	御 道 言 達 へ ず 仕 まつ り	あき	てしに思ふも 口 置しや	オク	置く		オク	動詞-非文語四段-連用形	ニ	ニ
26	つ	20_源氏物語01_04夕顔	80090	い ま、 な ど て 樂 り 添 ひ て	行か	てしに思ふも 口 置しや	イク	行く		イク	動詞-非文語四段-未然形	ニ	ニ
27	つ	22_源氏物語04_34若菜上	21190	ん ん こ と を 見 隠 し 教 へ	きこえ	てしに思ふも 口 置しや	キコエル	聞こえる		キコユ	動詞-非文語下二形-連用形	ニ	ニ
28	つ	22_源氏物語04_34若菜上	22340	権 中 納言 の 朝 臣 の 位 に	あり	てしに思ふも 口 置しや	アル	有る		アリ	動詞-非文語上二形-連用形	ニ	ニ
29	つ	20_源氏物語01_04夕顔	12410	り め た も かれ に ほ の ほ	見	てしに思ふも 口 置しや	ミル	見る		ミル	動詞-非文語上二形-連用形	ニ	ニ
30	つ	22_源氏物語04_34若菜上	31340	か な ら ず に け 申 さ せ	たまひ	てしに思ふも 口 置しや	タマフ	給う	尊敬	タマフ	動詞-非文語四段-連用形	ニ	ニ
31	つ	22_源氏物語04_34若菜上	32830	世 の 甲 を 御 心 と 言 ひ	たまひ	てしに思ふも 口 置しや	タマフ	給う	尊敬	タマフ	動詞-非文語四段-連用形	ニ	ニ
32	つ	22_源氏物語04_34若菜上	45090	せ たまへ し 御 氣 色 を	見	てしに思ふも 口 置しや	ミル	見る		ミル	動詞-非文語上二形-連用形	ニ	ニ
33	つ	22_源氏物語04_34若菜上	45040	た まへ り し 御 氣 色 を	見	てしに思ふも 口 置しや	タマツル	奉る		タマツル	動詞-非文語四段-連用形	ニ	ニ
34	つ	22_源氏物語04_34若菜上	45340	あ ら じ け し と 心 と き め し	し	てしに思ふも 口 置しや	ス	為る		ス	動詞-非文語上二形-連用形	ニ	ニ
35	つ	22_源氏物語04_34若菜上	49210	か く と り わ ぎ て 聞 き あ き	たてまつり	てしに思ふも 口 置しや	タマツル	奉る		タマツル	動詞-非文語四段-連用形	ニ	ニ
36	つ	22_源氏物語04_34若菜上	68780	中 納言 の 位 に	あり	てしに思ふも 口 置しや	アル	有る		アル	動詞-非文語上二形-連用形	ニ	ニ

図6 助動詞「つ」の上接動詞の検索結果(一部)

このデータをピボットテーブルで集計することで、助動詞別に高頻度な上接動詞をリストアップすることができる(図7)。

助動詞	つ	助動詞	ぬ
行ラベル	データの個数 / キ	行ラベル	データの個数
給う	284	成る	726
侍る	123	給う	725
見る	116	侍る	245
為る	107	有る	137
思う	103	出でる	134
有る	90	過ぎる	134
聞こえる	83	止む	108
奉る	70	経る	95
過ぎす	46	失せる	86
言う	44	入る	84
見える	38	果てる	83
思す	31	返る	72
初める	29	来る	65
遣る	27	為る	61
捨てる	26	参る	60
申す	25	立つ	57
宣つ	23	渡る	55
果てる	22	暮れる	52
聞く	21	忘れる	44
成す	21	消える	42
来る	19	絶える	41
参る	19	更ける	39
置く	19	知る	38
取る	19	明する	38
許す	18	寝る	38
得る	15	出で来る	37
侍ふ	13	罷出づ	35
渡る	13	泣く	33
変える	13	亡くなる	32
止す	13	降る	32
留める	13	隠れる	32
渡す	12	思う	31
覚える	12	散る	31
下ろす	12	止まる	29
おわします	12	思す	27
明かす	12	居る	26
暮らす	11	別れる	26
ものす	11	おわす	26
扱げる	10	おわします	26
失う	10	上る	26
おわす	10	見える	22

図7 助動詞「つ」「ぬ」の高頻度の上接動詞(一部)

## 7. おわりに

検索例で見たように、従来であれば大変な労力と時間を要していた検索・集計作業を、極めて簡単に行うことが可能になった。また、従来では不可能であった高度な組み合わせ検索が可能になった。今後、通時コーパス用『中納言』を用いて、単に研究を省力化するだけでなく、これまで不可能であった新次元の古典語研究がなされ、有益な研究成果が生み出されることに期待したい。

現在、通時コーパス用『中納言』の公開範囲は共同研究者の一部に限定しているが、今後は公開範囲を拡大していく予定である。

## 文 献

- 小木曾智信・中村壮範・鈴木泰山・八木豊・山崎誠・前川喜久雄(2011)「コーパス検索システム「中納言」デモンストレーション」『日本語コーパス完成記念講演会予稿集』pp.43-46
- 小木曾智信ほか(2012)『和文系資料を対象とした形態素解析辞書の開発』科研費 基盤研究 (C)「和文系資料を対象とした形態素解析辞書の開発」(課題番号 21520492) 研究成果報告書 ([http://dl.dropbox.com/u/73297026/report/unidic-EMJ\\_report2012.pdf](http://dl.dropbox.com/u/73297026/report/unidic-EMJ_report2012.pdf) からダウンロード可能)
- 小椋秀樹・須永哲矢(2012)『中古和文 UniDic 短単位規程集』科研費 基盤研究 (C)「和文系資料を対象とした形態素解析辞書の開発」(課題番号 21520492) 研究成果報告書 2 ([http://dl.dropbox.com/u/73297026/report/unidic-EMJ\\_rulebook2012.pdf](http://dl.dropbox.com/u/73297026/report/unidic-EMJ_rulebook2012.pdf) からダウンロード可能)
- Toshinobu Ogiso, Mamoru Komachi, Yasuharu Den and Yuji Matsumoto. (2012) UniDic for Early Middle Japanese: a Dictionary for Morphological Analysis of Classical Japanese. In *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC)*, pp.911-915. Istanbul, May 2012. ([http://www.lrec-conf.org/proceedings/lrec2012/pdf/906\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2012/pdf/906_Paper.pdf) からダウンロード可能)

## 関連 URL

- コーパス検索アプリケーション『中納言』(BCCWJ) <http://chunagon.ninjal.ac.jp/>
- NINJAL 通時コーパスプロジェクト ホームページ <http://www.historicalcorpus.jp/>
- 中古和文 UniDic <http://www2.ninjal.ac.jp/lrc/index.php?UniDic>