

国立国語研究所「日本語研究・日本語教育文献データベース」 の特長—「専門特化型」データベースの有用性—

中野真樹 渡辺由貴 早田美智子 横山詔一
国立国語研究所研究情報資料センター

今日、先行研究の検索・参照等のために、様々なリファレンスデータベースが作成されている。国立国語研究所は2010年に「日本語研究・日本語教育文献データベース」を公開した。このデータベースは日本語学・日本語教育研究の文献に特化している。このような特定の専門分野の文献にしばって作られている「専門特化型」データベースが、独自の観点から情報の収集・選択・整理を行っているという特性を生かし、多分野にわたる文献をナビゲートしている網羅的なデータベースとともに活用されることが、リファレンスデータベース全体、また、各学界の進展に寄与すると期待される。

The Features of the “Bibliographic Database of Japanese Language Research” by the National Institute for Japanese Language and Linguistics:

On the Utility of Specialized Databases

Maki NAKANO Yuki WATANABE Michiko HAYATA Shoichi YOKOYAMA
Center for Research Resources
The National Institute for Japanese Language and Linguistics

Currently, various reference databases are being constructed and utilized to search for and refer to published research. In 2010, the National Institute for Japanese Language and Linguistics went online with its "Bibliographic Database of Japanese Language Research," which covers literature in the fields of Japanese language and Japanese-language teaching. Such databases, focusing on particular fields and incorporating information carefully selected by experts, complement the comprehensive databases that cover a wide range of academic disciplines, and should play a significant role in the future, both in the specific development of reference databases and in academic endeavors in general.

1. まえがき

先行研究の検索・参照等のためには、様々な方法があるが、現在ではインターネットにより、膨大な学術情報が入手可能となっているとよい。特に、研究の要となる雑誌論文を中心とした文献については、その書誌事項を検索するための様々なリファレンスデータベースが存在している¹。

雑誌論文の検索の際には、例えば医学・生物等の分野では「PubMed²（米国立医学図書館）」等を利用すれば、国内外の雑誌論文を検索することができる。日本国内の雑誌論文については、多くの情報提供機関と提携し、多分野にわたる文献をナビゲートしている「CiNii³（国立情報学研究所）」等で検索できる。人文科学の分野では、日本の雑誌論文を検索する際には、特に「CiNii」が使われることが多い。「CiNii」は採録件数も多く、様々な機能をそなえたデータベースではあるが、一方、特定の研究分野にしばったリファレンスデータベースも存在している。国立国語研究所が提供する「日本語研究・日本語教育文献データベース」（以下本データベースとする）もそれにあたる。このような専門分野に限定する方向のデータベースが、独自の観点から情報の収集・選択・整理を行っているという特性を生かして発展し、活用されることが、リファレンスデータベース全体、また、各学界の進展に寄与すると期待される。

¹ 雑誌論文以外の学術情報検索の方法としては、書籍情報を検索できる「WebCat」（<http://webcat.nii.ac.jp/>）、「NDL-OPAC」（<http://opac.ndl.go.jp/>）、「インターネット書籍販売サイト「Amazon」（<http://www.amazon.co.jp/>）、「学術資料全般を検索できる「Google Scholar」（<http://scholar.google.co.jp/>）等がある。

² <http://www.ncbi.nlm.nih.gov/pubmed/>

³ <http://ci.nii.ac.jp/>

本発表では、まず国立国語研究所「日本語研究・日本語教育文献データベース」(<http://www.ninjal.ac.jp/database/bunken/>)の概要を示す。そして、本データベースをリファレンスデータベースの中の「専門特化型」データベースとして位置づけ、その特長を述べながら、その特性を生かす方向性を模索することで、リファレンスデータベースの展望を考える上での一助としたい。

2. 「日本語研究・日本語教育文献データベース」の概要

本データベースは、国立国語研究所が作成・提供する、日本語学・日本語教育に関する雑誌論文情報検索のためのリファレンスデータベースである。2011年1月に公開された。

【図1 本データベースのなりたち】



本データベースは、書籍として論文情報を提供してきた『国語年鑑』と『日本語教育年鑑』のデータを統合したものであり、掲載件数は約17万2千件である。今後、年3回程度新規データの追加更新を行っていく予定である。

所収データは、『国語年鑑』が1950年以降、『日本語教育年鑑』が1960年以降に発表された、約60年分の論文の情報である。採録された論文原本は、国立国語研究所の研究図書室に配架されている。

『国語年鑑』は日本語学に関する文献情報、『日本語教育年鑑』は日本語教育に関する文献情報を収録したものであったが、両者は関わりが深く、両方の分野に携わる研究者も多い。

両年鑑のデータを統合するにあたっては、重複データを統合・整理するのに加え、公開する項目を追加するなどの充実をはかっている。

項目の構成は下記の通りである。

1. 文献番号
2. 研究図書室請求番号
3. 著者名
4. 著者名別表記
5. 論文名
6. 論文名別表記
7. 誌名
8. 巻号
9. ページ
10. 発行機関
11. 発行年月
12. キーワード
13. 章タイトル
14. 分野

【図2 項目構成（検索画面）】

日本語研究・日本語教育文献 検索

「日本語研究・日本語教育文献」に関する論文を検索できます。

結果表示件数 / ページ:	<input checked="" type="radio"/> 100件 <input type="radio"/> 50件 <input type="radio"/> 20件	
ソートモード:	<input checked="" type="radio"/> 表示するデータでソート <input type="radio"/> 検索結果全体でソート	
著者名 [?] <input type="text"/>	すべて ▾	
論文名 [?] <input type="text"/>	すべて ▾	
誌名 [?] <input type="text"/>	すべて ▾	
発行機関名 [?] <input type="text"/>	すべて ▾	
発行年 [?] <input type="text"/> ~ <input type="text"/>		
キーワード・章タイトル [?] <input type="text"/>	すべて ▾	
分野 [?] <input type="checkbox"/>	<input type="checkbox"/> 日本語学一般 <input type="checkbox"/> 日本語史 <input type="checkbox"/> 音声・音韻 <input type="checkbox"/> 文字・表記 <input type="checkbox"/> 語彙・用語 <input type="checkbox"/> 文法 <input type="checkbox"/> 待遇表現(注) <input type="checkbox"/> 文章・文体 <input type="checkbox"/> 古典の注釈 <input type="checkbox"/> 方言・民俗 <input type="checkbox"/> 日本語情報処理 <input type="checkbox"/> コミュニケーション <input type="checkbox"/> マスコミュニケーション <input type="checkbox"/> 国語問題・言語問題 <input type="checkbox"/> 国語教育 <input type="checkbox"/> 日本語教育 <input type="checkbox"/> 言語学 <input type="checkbox"/> 資料 <input type="checkbox"/> 書評・紹介 <input type="checkbox"/> 索引 <input type="checkbox"/> 辞典・用語集	

● 著者名、論文名、誌名、キーワード・章タイトルの項目にその語を含んでいるデータを表示します。

統合にあたって、特に以下の項目について整備した。

(2.研究図書室請求番号)

閲覧の利便性のために研究図書室の請求番号を付与した。

なお、採録されたデータの論文原本は、国立国語研究所の研究図書室にある。ここには、日本語学・日本語教育関連を中心とした文献が収められており、また、この分野の学会・研究会とのつながりが強く、寄贈される文献も多い。

(13.章タイトル)

論文名やキーワードからは拾いきれない情報を付与することで、検索の精度を向上させるとともに、その論文のおおまかな内容の把握を可能にした。

(14.分野)

一定の基準をもとに付与したメタデータである。「文法」「方言」「日本語情報処理」等があり、検索の際の参考情報となる。動向分析にも利用できる。

以下は、検索結果の一覧画面である。

【図3 検索結果一覧画面】

選択したデータをテキスト(CSV)に書き出す(ページごと) 文字コード: UTF-8

- データ一覧の上下に表示されている項目名をクリックすると、その項目の昇順降順(文字コード順)で並べ替えることができます。
- ここにチェックを入れ、「表示データをテキストに書き出す」ボタンをクリックすると、選択したデータのみ書き出します。データをExcelで開く場合「外部データの取り込み」を使用してください。
※詳細についてはMicrosoftのWebサイトを参照してください。
<http://office.microsoft.com/ja-jp/excel-help/HP010101234.aspx?CTT=1>
- 直接Excelで開くとデータが数個誤変換されることがありますのでご注意ください(特に項目「巻号」)。
- 左端の番号をクリックすると、「詳細表示画面」が表示され、すべての情報を見ることができます。印刷までできますが、この形ではテキストに書き出すことはできません。

No.	一括選択	著者名	論文名	誌名	巻号	ページ	発行機関	発行年月	キーワード	章タイトル
1	<input type="checkbox"/>	三井 はるみ	条件表現の地理的変異—方言文法の体系と多様性をめぐって—	日本語科学		pp.143-164	国立国語研究所	2009.04	方言文法全国地図、順接仮定条件表現、体系の多様性、津軽方言の「ら」、佐賀方言の「キー」	1『方言文法全国地図』と方言文法研究 2取り扱う表現分類 順接仮定条件表現形式の全国的地理差 4.春森県津軽方言 接仮定条件表現 5.佐賀方言の順接仮定条件表現 6.方言 接仮定条件表現に見られる傾向 7.方言の文法体系の多様 記述に向けて
2	<input type="checkbox"/>	車田 秀也	理解の問題と発話産出の問題— 理解チェック連続における「うん」と 「そう」—	日本語科学		pp.43-66	国立国語研究所	2009.04	修訂、理解チェック、訂正、言葉探し、認識	1はじめに 2.研究のコンテキスト 3.理解チェックとその確認 行的訂正・言葉の試行的提示とその受け入れ 5.認識チェックの確認 6.独自の貢献を見いだすこと 7.結論
3	<input type="checkbox"/>	向山 陽子	第二言語習得において学習者の適性が学習成果に与える影響— 言語分析能力・音韻的短期記憶・ワーキングメモリに焦点を当てて—	日本語科学		pp.67-90	国立国語研究所	2009.04	言語適性、言語分析能力、音韻的短期記憶、ワーキングメモリ	1はじめに 2.先行研究 3.研究方法 4.結果 5.考察 6.まとめ 今後の課題
4	<input type="checkbox"/>	宮岡 弥生;玉岡 真津雄;林 炫甫;池 映任	韓国語を母語とする日本語学習者による漢字の書き取りに関する研究— 学習者の語彙力と漢字が含まれる単語の使用頻度の影響—	日本語科学		pp.119-130	国立国語研究所	2009.04	漢字、韓国語母語話者、日本語学習者、語彙使用頻度、音韻的類似性	1はじめに 2.漢字文化圏と韓国語 3.本研究の目的 4.テスト概要 5.日本語の語彙力と語彙使用頻度に関する分析と考 1日韓の音韻的類似性に関する分析と考察 7.考察
5	<input type="checkbox"/>	永田 良太	複文発話の構文的特徴と聞き手の言語的反応との関係— グラド、タラ、カマを中心として—	日本語科学		pp.5-22	国立国語研究所	2009.04	從節、主節、文の層層構造、あいつち	1はじめに 2.発話に対する聞き手の反応 3.分析資料と考 4.從節末と主節末における聞き手の言語的反応の比較 : 助詞別の比較 6.まとめと今後の課題
6	<input type="checkbox"/>	田野村 忠温	サ変動詞の活用ゆれについて— 統一大規模な電子資料の利用による分析の精査—	日本語科学		pp.91-103	国立国語研究所	2009.04	サ変動詞、活用ゆれ、大規模な電子資料、国会会議録、ウェブサーバ	1はじめに 2.国会会議録データに基づく分析 3.ウェブサーバ 基づく分析 4.おわりに

検索結果の一覧画面では、おおまかな内容の把握ができるようになっている。さらに、項目名をクリックすることで簡単にソートができ、項目ごとにまとまった情報も得られる。一覧表示画面(図3)一番左の「No.」欄にある数字をクリックするとデータの詳細画面になり、そのデータのより詳細な情報を見ることができる。また、選択したデータをダウンロードすることもできる。

3. 人文科学分野からみたりファレンスデータベースと「日本語研究・日本語教育文献データベース」の位置づけ

以下に、人文科学分野における雑誌論文検索のためのリファレンスデータベースについて述べる。

研究においては、先行論文の参照が必須である。そして、被引用件数が論文の質を評価する一つの指標ともなっている。学術雑誌の重要度については、インパクトファクター¹という指標が有効な研究分野もある。

しかし、人文科学分野では事情が異なり、例えば日本語学のような、日本語で執筆されることが主であり、かつ主要な雑誌であっても基本的にこの指標が使いにくい領域では、論文の重要度が客観的に測りにくい。また、日本語学の分野では、雑誌種別が「大学紀要」とされるような論文もよく参照

¹ Impact factor (文献引用影響率) 特定の1年間において、ある特定雑誌に掲載された「平均的な論文」がどれくらい頻繁に引用されているかを示す尺度。

< <http://science.thomsonreuters.jp/products/jcr/support/> > (アクセス日: 2011/11/7)

されている。そのため、その領域内の情報のみをくまなく収集したリファレンスデータベースも必要になる。

本発表では人文科学分野でよく使われているリファレンスデータベースについて、データの掲載範囲の観点から、多くの研究分野を網羅的に採録している「網羅的」データベースと、特定の研究分野を対象として文献を採録している「専門特化型」データベースに分類して考察する。国内で利用される人文科学分野の「網羅的」データベースとしては、「CiNii」(国立情報学研究所)、「雑誌記事索引検索¹」(国立国会図書館)や「JDream II²」(科学技術振興機構)等が挙げられ、「専門特化型」データベースとしては、「国文学論文目録データベース³」(国文学研究資料館)や「東洋学文献類目検索⁴」(京都大学人文科学研究所)等が挙げられる。本データベースも、「専門特化型」データベースに位置づけられる。なお、「網羅的」データベースとはいえ、もちろん、全世界のすべての資料を網羅することは物理的に難しいであろう。たとえば雑誌論文という媒体に特化したもの、国内で出された文献に特化したもの、という形でなんらかに特化することにはなる。しかし、ここでは、研究分野には限定せず網羅的に文献を採取しているということで「網羅的」データベースと位置付けておく。

【図4 「網羅的」データベースと「専門特化型」データベース】



次に、人文科学分野で利用されるデータベースをいくつか挙げ、その特徴を採録件数の多い順に【図5】にまとめた(2011年10月現在、当該ページで得られた情報による)。

【図5 上記リファレンスデータベースの特徴⁵】

	データベース名	研究分野による採録範囲の限定	件数	情報源	一次情報へのアクセス
「網羅的」データベース	JDREAM II	—	約5600万件	他機関からの情報	一部有
	CiNii	—	約1200万件	他機関からの情報	一部有
	雑誌記事索引検索	—	約1036万件	独自(国会図書館)	なし
「専門特化型」データベース	国文学論文目録データベース	有(国文学)	約48万6千件	独自(国文学研究資料館)	なし
	教育研究論文索引検索	有(教育学)	約18万3千件	独自(国立教育政策研究所)	CiNii経由
	日本語研究・日本語教育文献データベース	有(日本語学・日本語教育)	約17万2千件	独自(国立国語研究所)	なし(計画中)
	東洋学文献類目検索	有(東洋学)	不明	独自(京都大学)	なし

¹ <http://opac.ndl.go.jp/Process>

² <http://pr.jst.go.jp/jdream2/>

³ <http://base1.nijl.ac.jp/~ronbun/>

⁴ <http://ruimoku.zinbun.kyoto-u.ac.jp/ruimoku/>

⁵ データベースによっては、雑誌論文以外のものも件数に含んでいる場合がある。

「網羅的」データベースと「専門特化型」データベースに大きく二分したが、採録範囲をどこまでとするか、その線引きはそのままそれぞれのデータベースの特長ともなっている。

データベースの規模は数千万単位に及ぶものから、数十万規模のもの、さらには【図5】には掲載していないがごく小規模なものまで存在する。採録を特定の範囲にしばらない「網羅的」データベースは千万件台になっているが、採録範囲を限定すればその分件数は減る。本データベースは「国文学論文目録データベース」や「教育研究論文索引検索」等とならび、ある研究分野範囲のデータを採録対象としている。結果としてこれらのデータベースは十万件台の規模となっている。

また、採録範囲を限定した「専門特化型」データベースは、独自の情報源からデータをつくって持っているところが多いようである。例えば、前述のとおり、本データベースでも国立国語研究所の図書室から情報の提供を受けている。「網羅的」データベースは、その性質上、他機関から得た情報をナビゲートする形をとることになるが（「CiNii」等）、国会図書館のように、独自の情報源を持つ形をとっているものもある（「雑誌記事索引検索」）。

なお、基本的には文献の書誌事項等、二次情報のみを収録するリファレンスデータベースが多いが、一部、本文とのリンクが設けられることで全文データベースの側面を持ち合わせるものも存在する。

以上のように、人文科学系で使われるデータベースの特長を概観したが、そのなかでもやはり、論文検索の際、採録件数が多く、機能が充実している「網羅的」データベースはよく利用されている。しかし、リファレンスデータベースが一つの「網羅的」データベースに収斂されていくという方向で十分とは言えないだろう。各研究分野で「専門特化型」データベースが整備され、「網羅的」データベースとともにあり、目的に応じて使い分けられることがユーザーにとって望ましいといえるのではないだろうか。

4. 「専門特化型」データベースとしての本データベースの特長

ここでは、本データベースの特長を整理することによって、「専門特化型」データベースの有用性は何かということについて考えてみたい。

(1)独自の情報源

本データベースは、情報源に日本語に関する専門図書館である国立国語研究所研究図書室をもっており、所蔵する原論文から直接書誌データを入力している。この分野の学会・研究会とのつながりが強く、日本語学および日本語教育に関する研究文献が多く寄贈されている。なお、本データベースに関しては、国立国語研究所が事業として継続的に雑誌を集めて整理してきたので、過去60年程度の論文情報の蓄積もあり（2. 参照）、日本語学研究では重要となる長期的な先行研究の概観に一役買っている。茂木(2010)[1]によると、『国語年鑑2008年版』掲載の2007年発行の大学紀要類に関して言えば、約9%の雑誌は「CiNii」に当該巻号の情報がないとのことである¹。これに加え、国立国語研究所には、紀要以外にも広く流通していない文献等が寄贈されることもあり、海外の文献に関しても、日本語学に関連するものを収集している。また、上記のような60年分のデータの蓄積も含めると、「網羅的」データベースに採録されていない雑誌論文の件数はさらに多いものと考えられる。

(2)専門家による取捨選択と整理

3. で述べたとおり、この研究分野ではインパクトファクターのような客観的な指標が通用しないため、専門分野に関する論文は、掲載雑誌に関わらずくまなく収集しておく必要がある。研究分野に関わる論文のみが掲載される専門誌の他に、様々な研究分野の論文が掲載される大学紀要等に収録された論文から、日本語学・日本語教育に関する論文のみを抽出する必要がある。そのため、専門家が原論文にあたって日本語学・日本語教育研究の観点から収録するものを選択し、同時に研究内容に関するメタデータ（分野）も付与している。

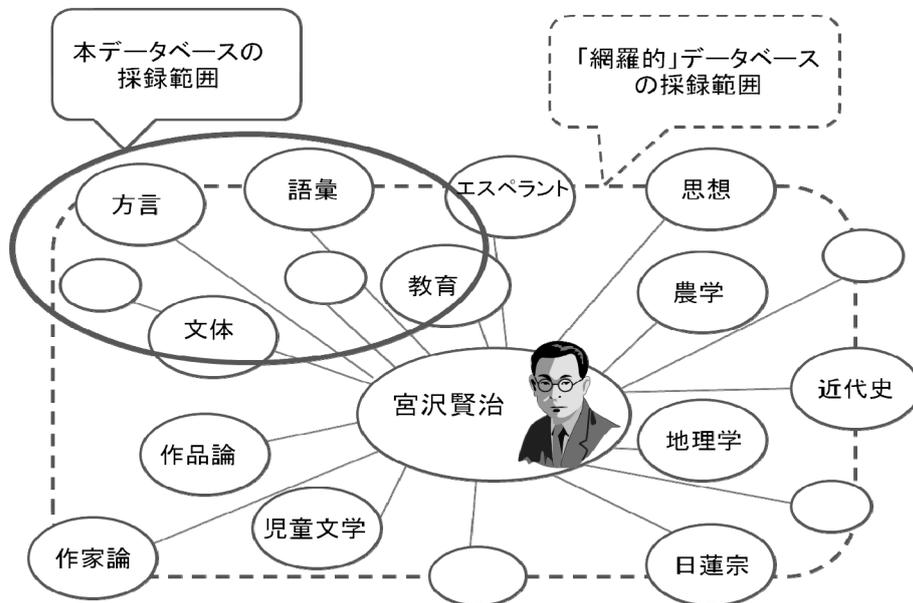
(3)検索の利便性

本データベースには、日本語学・日本語教育に関する文献が選択され、採録されている。論文のタイトルにも用いられやすく、「網羅的」データベースでは検索結果が膨大になってしまう研究論文を

¹ 他の研究分野でも「CiNii」との重複率は興味を持たれており、江草・高久(2011)[2]によると、「教育研究論文索引データベース」の収録論文のうち、重複データは約6割であるとのことである。

検索する際等には、当該研究分野の論文の中から効率的に探すことができる「専門特化型」データベースが有用である。また、付与されたメタデータ（分野）を検索の際の参考情報として使うこともできる。

【図6 宮沢賢治研究のイメージ図】



以下、宮沢賢治を例に具体的に説明する。「CiNii」にて論文タイトルに「賢治」と入力して検索すると、3499件（2011年10月27日現在）ヒットした。「国文学論文目録データベース」では2719件、本データベースは100件であった。

「CiNii」は網羅的に様々なジャンルの論文を採録しているのだから必然的に件数が多くなる。「国文学論文目録データベース」は、国文学の分野の研究情報を採録した「専門特化型」データベースであり、宮沢賢治の研究についてもコアとなるデータベースなので、データ数も多くなっている（論文集等、雑誌以外の媒体からもデータを採録していることもあるので、検索の仕方によっては「CiNii」より多くの件数がヒットすることもある）。これに対し、本データベースは、件数としてはそれほど多くない。しかし、宮沢賢治研究のうち、オノマトペや文体など、言葉の観点から書かれた論文を採録しているのだから、「宮沢賢治の作品を日本語学的観点から研究したい」等と漠然と研究テーマを探しているような場合、本データベースであたりをつけてみるという使い方ができる。

本データベースではさらに、「語彙・用語」「文章・文体」等の分野情報が付与されており、それらを利用して内容をしぼりこむこともできる。現状では、「網羅的」データベースでは、キーワード検索等を利用して、タームレベルのしぼりこみをするのは可能である。しかし、「日本語学」「日本語教育」のような研究分野を示す語は、論文タイトルやキーワードには出てきにくいから、ある研究分野の範囲で論文をしぼりこむことは難しい。特にリファレンスデータベースの検索に不慣れた初学者や、あらたに学際的研究を行う際や、関係するジャンルの研究成果を概観する場合にはおおいに参考となる。

このように、既に特定の研究分野でしぼりこまれていることが有益な場合もある。また、採録対象の雑誌の関係で、他のデータベースには掲載されていないデータが見られることもあるので、「専門特化型」データベースを確認しておくことは意味があるだろう。

(4) 目録としての役目

本データベースでは、2.でも述べたとおり、必要な文献の一覧をデータでダウンロードできる。また、年ごと・分野ごと等で検索してカスタマイズしておく、自分に必要な分野の研究目録一覧を作成することもできる。例えば、「2010年に発表された方言研究」の論文一覧を作る等の利用方法もある。かつて『国語年鑑』『日本語教育年鑑』にあった文献一覧にあたるリストや、さらに加工して、他のシステムで読み込むための業績一覧用のデータを作成する等、各自の目的にあわせて利用できる。

【図7 検索結果一覧ダウンロード画面(CSV)】

▼	A	B	C	D	E	F	G	H	I	J	K
1	文献番号	研究図書室 請求番号	著者名	論文名	誌名	巻号	ページ	発行機関	発行年月	キーワード	章タイトル
2	1000001	092 N77 25	永田 良大	複文発話の構文的特徴と聞き手の言語的反応との関わり—クワド、タラ、カラを中心に—	日本語科学	25	pp.5-22	国立国語研究所	2009.04	従属節、主節、文の階層構造、あいづち	1.はじめに 2.発話に対する聞き手の反応 3.分料と考察対象 4.従属節末と主節末における階層の言語的反応の比較 5.接続助詞の比較 6.
3	1000003	092 N77 25	串田 秀也	理解の問題と発話産出の問題—理解チェック連鎖における「うん」と「そう」—	日本語科学	25	pp.43-66	国立国語研究所	2009.04	修復、理解チェック、訂正、言葉探し、認識	1.はじめに 2.研究のコンテキスト 3.理解チェックの確認 4.試行的訂正、言葉の試行的提示とそれ入れ 5.認識チェックとその確認 6.独自の貢献
4	1000004	092 N77 25	向山 陽子	第二言語習得において学習者の適性が学習成果に与える影響—言語分析能力、音韻的短期記憶、ワーキングメモリに焦点を当てて—	日本語科学	25	pp.67-90	国立国語研究所	2009.04	言語適性、言語分析能力、音韻的短期記憶、ワーキングメモリ	1.はじめに 2.先行研究 3.研究方法 4.結果 5.6.まとめと今後の課題
5	1000005	092 N77 25	田野村 忠温	サ変動詞の活用ゆれについて—統一大規模な電子資料の利用による分析の精密化—	日本語科学	25	pp.91-103	国立国語研究所	2009.04	サ変動詞、活用のゆれ、大規模な電子資料	1.はじめに 2.国国会議録データに基づく分析 3.コーパスに基づく分析 4.おわりに
6	1000007	092 N77 25	宮岡 弥生、玉岡 寛津、林 有博	韓国語を母語とする日本語学習者による漢字の書き取りに関する研究—学習者の語彙力と漢字が含まれる単語の使用頻度の影響—	日本語科学	25	pp.119-130	国立国語研究所	2009.04	漢字、韓国語母語話者、日本語学習者、語彙使用頻度、音韻	1.はじめに 2.漢字文化圏と韓国語 3.本研究の4.テストの概要 5.日本語の語彙力と語彙使用に関する分析と考察 6.日韓の音韻的類似性に関する分析 7.方言文法全国地図、方言文法全国地図、方言文法研究 2.取り分け表現分野 3.懸架仮定条件表現形式の全国的な差 4.青森県津軽方言の懸架仮定条件表現 5.方言の懸架仮定条件表現 6.方言の懸架仮定
7	1000009	092 N77 25	三井 はるみ	条件表現の地理的変異—方言文法の体系と多様性をめぐって—	日本語科学	25	pp.143-164	国立国語研究所	2009.04	条件表現、方言	1.はじめに 2.方言文法全国地図、方言文法研究 2.取り分け表現分野 3.懸架仮定条件表現形式の全国的な差 4.青森県津軽方言の懸架仮定条件表現 5.方言の懸架仮定条件表現 6.方言の懸架仮定

さらに、このように作った目録を、学界の動向分析等にも利用できる。山崎(1990)[3]や齋藤・新野(2002)[4]では、本データベースの元となった『国語年鑑』のデータに対する動向分析を行っている。『国語年鑑』でも2003年版から2009年版(2009年版はWEBでの公開)において、その年に出た日本語研究文献を数量的に扱い、分野ごとの比率をみたり、経年的な動きをみたりしている。今後は本データベースを利用して動向分析を行うことが可能である。

また、それと関連して、展望号を書くときの参考資料としても貢献できる。展望論文の執筆の際には、基本的にその分野の論文を全体的に見通す必要がある。しかし上述の通り、「網羅的」データベースでは日本語学だけ、日本語教育だけ、というように研究分野でしぼることは不可能であるので、「網羅的」データベースしか使えないと、執筆者にとっては大きな負担になる。例えば、膨大な情報の中での検索漏れの危険性や、不必要な情報の整理の煩雑さが考えられる。本データベースは既に専門家による選択が行われており、1年あたりの採録件数は最大でも4000件程度で、ノイズも少ないので、検索や一覧作成の際の負担は軽くなっている。実際、関連学会の展望号執筆の際にもデータが利用されており、国立国語研究所の研究図書室に原論文を確認にくる執筆者も多い。

このように、特定分野に限定したデータベースであるので、検索目的で使用するだけでなく、当該分野の研究一覧としての価値もある。

(5) 専門性を生かしたリンクの広がり

「専門特化型」データベースとして、上記のような特長を持つ本データベースは、その専門性を生かした発展の方向性が検討され、進行している。例えば、専門に特化したデータベースという視点から、その専門分野の研究に必要な他のデータベースとの横断的な検索を目指すことも可能である。

例えば、国立国語研究所では、単行本の論文集、学会の予稿集に掲載された文献のデータベース化およびこれらとの横断検索を検討中である。さらに、国立国語研究所が所属している人間文化研究機構の資源共有化システムとの横断検索も予定されている。

また、「網羅的」データベースにはない、学会との緊密な関係も強みである。現在、旧国語学会の学会誌をデータ化した「『国語学』全文テキストデータベース」掲載論文について、そのPDFおよびテキストと本データベースとにリンクを設ける計画も進んでいる。また、データ整備に関しても、ユーザーからの声が届きやすい。

5. 終わりに—リファレンスデータベースの将来とその可能性—

以上、本データベースの特長を述べてきたが、これは他の「専門特化型」データベースにも共通する点がある。

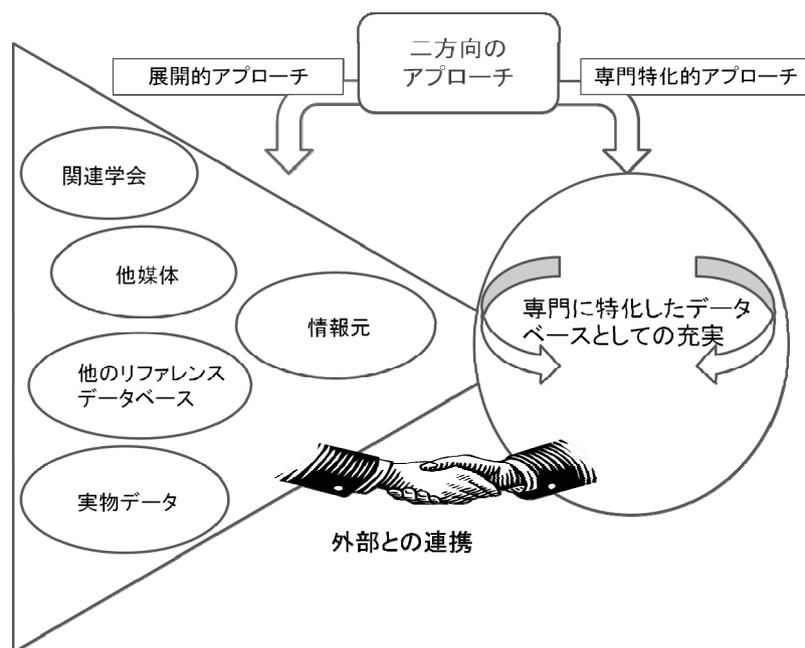
これをふまえると、「専門特化型」データベースの発展には、以下の二方向が考えられる。

一つは、取捨選択した情報を蓄積していくこと、ユーザーが使いやすいように整備していくことなど、専門に特化したデータベースとしての充実を図る方向である。

もう一つは、取捨選択したデータを基として、ユーザーによって有益な外部リンクや横断検索のシステムを設けること、関係学会と協力し、相互にデータ提供を行うことなど、外部との連携を広げていく方向である。

【図8】に「専門特化型」データベースの可能性についての概念図を示す。

【図8 「専門特化型」データベースの可能性】



「専門特化型」データベースは、専門に特化したデータベースとして充実させていく「専門特化的アプローチ」と、適切に取捨選択されたデータと外部情報とを結ぶ「展開的アプローチ」の二方向のアプローチをとることで、各学界の進展に貢献できるデータベースとなろう。また、「専門特化型」データベース、「網羅的」データベースに関わらず、個性豊かで良質なデータベースが共存することによって、データベース相互の情報提供や横断検索等も可能となり、リファレンスデータベース全体が活性化するだろう。

参考文献

- [1] 茂木俊伸(2010)「日本語研究論文情報の電子化の実態と論文検索スキル」『鳴門教育大学ジャーナル』7
- [2] 江草由佳・高久雅生(2011)「教育研究論文索引とCiNiiの重複率」『情報知識学会誌』21-2
- [3] 山崎誠(1990)「『日本語研究文献目録・雑誌編』にみる国語研究の動向」『国立国語研究所研究報告集』11
- [4] 斎藤達哉・新野直哉(2002)「『国語年鑑』に見る分野別文献数の動向—1985～2000年の雑誌掲載文献」『日本語科学』11