

国立国語研究所学術情報リポジトリ

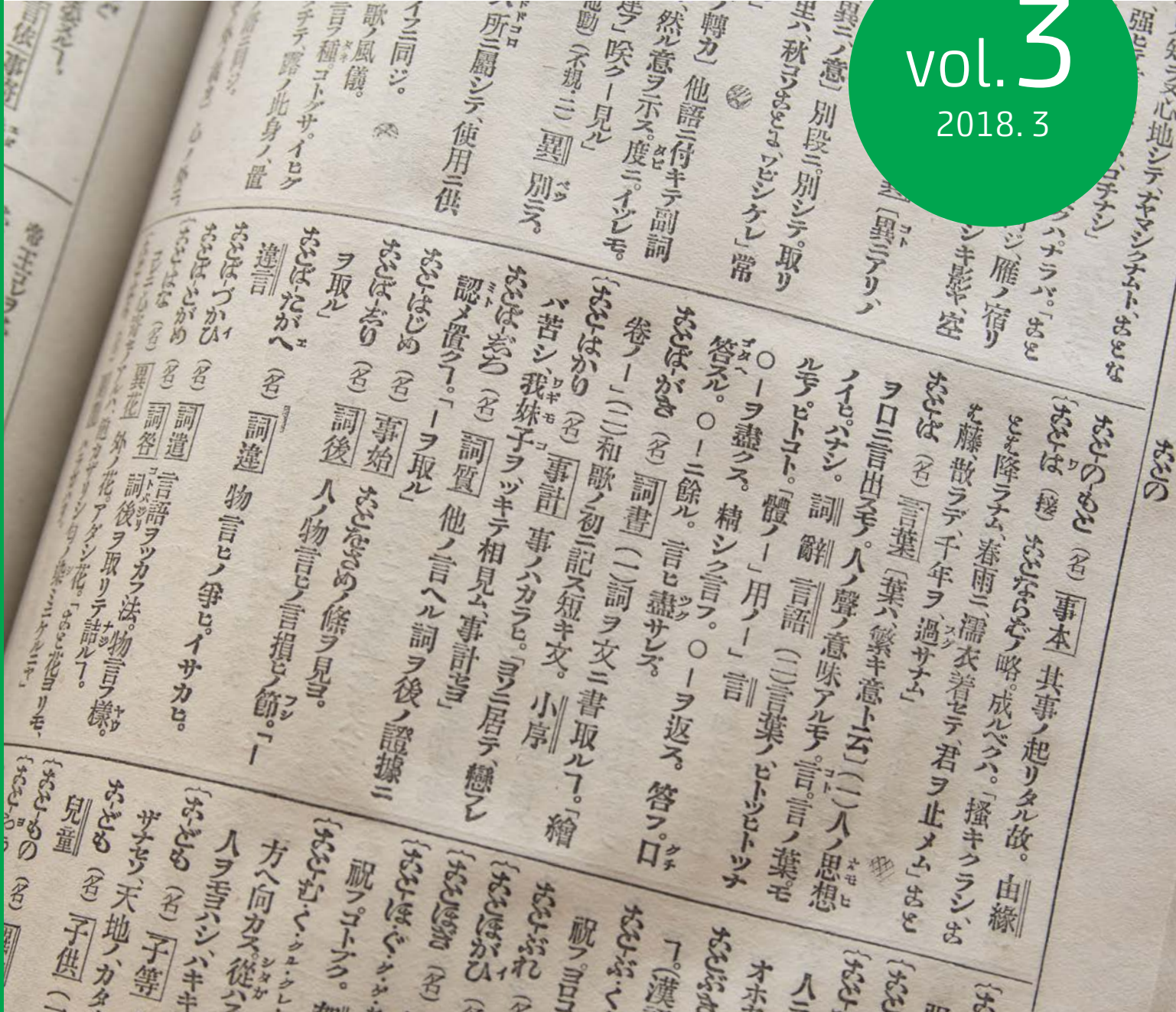
NINJAL Research Digest vol.3 (2018.3)

メタデータ	言語: jpn 出版者: 公開日: 2020-06-05 キーワード (Ja): キーワード (En): 作成者: 国立国語研究所研究情報誌編集委員会 メールアドレス: 所属:
URL	https://doi.org/10.15084/00002822

ことばの波止場

NINJAL Research Digest

vol.3
2018.3



特集

日本語の個性②

統語意味解析コーパス研究・日本語史研究・
音声言語研究

国立国語研究所・新所長から 田窪行則

研究者紹介 前川喜久雄 ビャーケ・フレスビッグ 蒙韞

著書紹介

ご挨拶にかえて

国立国語研究所長

田窪行則

TAKUBO Yukinori

専門領域は言語学、語用論、意味論、統語論、琉球諸語。九州大学文学部教授、京都大学文学研究科言語学専修教授などを経て、現職。



はじめに

平成29年10月1日付で大学共同利用機関法人人間文化研究機構国立国語研究所（以下国語研）所長に就任しました。

私は統語理論と呼ばれる文の組み立て方を理論的に研究することが専門ですが、日本語教育の経験があり、日本語教育を行っている間、自然言語処理の共同研究にも参加しました。最近は琉球諸語をはじめとする危機言語の調査も行っています。ここではそもそも私がどうして言語学を専門として選んだのか、なんで50代半ばを過ぎて琉球諸語の調査を始めたのかをお話して、自己紹介にかえたいと思います。

コトバへの目覚め

私は中学校のころは漠然と建築家やインテリアデザイナーになろうと思って絵を描いていました。高校に入って、美術部に入っすこし本格的に絵を始めたのですが、すぐに才能がないことがわかって、美術部をやめ、(多少不純な理由で)英会話クラブにかかりました。

そのようにして英語を勉強してい

たのですが、市立図書館で英作文の本を探していた時に偶然金田一京助の「北の人」という随筆に出会います。これは彼が樺太アイヌの調査に行ったときの話です。彼は調査に行き詰まり、調査の継続をあきらめようかと、無聊に絵を描いていました。そのとき周りにいた子供たちが寄ってきて、絵の中のものを指さしながら、その名前を樺太アイヌ語で言います。そこでひらめいた金田一は画用紙にでたらめの線を書きます。そうすると子供たちは口々に「ヘマタ」と叫びます。すでに知っていた北海道アイヌ語との類推から「ヘマタ」は「何？」に当たることばであると確信した金田一は次々にものを指しながらヘマタと言っているいろいろな単語の形を調べます。「ヘマタ」という魔法の言葉を得た金田一は一か月ほどで、樺太アイヌ語の文法概要と語彙集、3000行の叙事詩の採録を終えて帰任します。この随筆を読んだときの感動はいまでも忘れません。

大学入学の頃

ただ、このころは言語学という学問の存在をちゃんと知っていたかどうかわかりません。言葉は好きだっ

たようで、国語の時間にならった文節の係り受けに凝って、係り受けの構造図を毎日書いたり、それを英語の文の構造分析に使ったりしていました。発音記号が好きで、発音記号を見ながら、何回も発音練習してそこそこきれいな英語が話せるようになりました。そうこうしているうちに大学受験になるのですが、多分英語教師になるつもりで、大阪外国語大学（現在の大阪大学）に入ろうと思ったのだと記憶しています。当時は一期、二期と大学が分かれていて、大阪外大は二期校だったので、一期校をどこか受けようと思って京都大学を受けたのだと思います。苦手科目だった数学がその年は非常にやさしかったため、他の受験生と差がつかず、京都大学に受かってしまったので、大阪外大は受けませんでした。受けて合格していれば影山前所長の一年後輩になって同じ授業を受けていたかもしれません。

1969年大学紛争の真っ最中に京都大学に入って、一年間ほとんど授業がなかったので、時々英会話サークルや体操部の練習に出かけながら、毎日英語の小説を読んでいました。たぶん最初は子供向けの優しい本や、外国人向けにやさしく書き直した小

説を読んでいたのだと思うのですが、だんだんと普通の小説を読むようになって、それほど問題なく英語が読めるようになりました。話したり聞いたりとはもともと毎日英語の放送を聞いていたので問題はありませんでした。デンマークの言語学者のイエスペルセン (Jespersen) の書いた英文法の本や、アメリカの構造主義言語学者で構造主義言語学を英語教育に応用したフリーズ (Fries) の書いた構造主義の入門書なんかを読んで楽しかったので、英語学を専攻しようと思ったのですが、当時の京都大学の英語・英文科は英語学の先生がおらず、英語学を専攻する学生は受け入れてくれませんでした。

理論研究への道

入学して一年たって授業が再開し、担任の先生が大橋保夫という有名なフランス語学者だったり、渡辺実、阪倉篤義というこれも有名な国語学の先生の開講されていた言学という言語学の授業に感動したりして、結局言語学を専攻することにして、言語学講座に入りました。当時の言語学講座は西田龍雄先生というえらい先生がいたのですが、私が入ったときはロシアに長期出張に行っていて誰もおらず、東京外大から徳永康元という先生が来て集中講義をしてくれました。なぜかその先生の授業には私一人しか出席しておらず、二人きりで教えてもらいました。

その後、西田先生が帰国し、授業が本格的に始まりました。言語学講座でいろんな言語を勉強しながら、構造主義言語学、生成文法、日本語、朝鮮語、満洲語（清朝の言語）の文献学などを専門的に学びましたが、文献学は、記憶力がすぐれ、また文献そのものが好きでないとできないのですぐに自分には才能がないこと

がわかりました。言語学講座では、金田一のように外国の奥地に出かけ、現地で収集したデータで文法や辞書を作るというフィールドワークをやる人もいました。その苦労語を聞くとともに自分にはできないと思い、フィールドワークも諦めて、自分にもできそうな統語理論をやりました。

日本語教育の経験も

大学院の博士課程を出るときに韓国の大学で日本語を教える職があり、韓国の慶州というところで2年間現代日本語や古文を教えました。そのあと運よく神戸大学の教養部の職を得て、日本語・日本事情の教師をやったのですが、日本の大学で日本語を教えるとなると、共通の言語はないわけで、日本語のみで日本語を教えるため、ずいぶん苦勞しました。このころ、日本語教育の必要性から、統語理論的な研究より、日本語の記述に関心を移し、日本語教育に資するような文法書や留学生用の練習帳などを作りました。また、日本語教育用の文法書が実は言語処理に利用可能ということもあり、工学的な言語処理関係の方との接点も増えていきました。

宮古語への関心～研究の原点へ

九州大学の言語学講座に移ってからもちまた、2000年に京都大学に移ってからもしばらく理論的な研究や教育を行ってきたのですが、2006年1月全くの偶然でUCLAの岩崎勝一さん、カナダアルバータ大学の犬野剛さんたちと宮古島の池間方言の調査を始めます。これは京都大学の総長裁量経費をいただいて学生とのフィールドワークに使えると言われたためです。それまで敬して遠ざけていたフィールドワーク調査に出かけ

ました(写真)。池間方言は琉球諸語のひとつの宮古語の方言ですが、日本語の共通語とは全く異なり別の言語です。それが調査によってすこしずつ理解できるようになっていきます。すこし経つと全く理解できなかった談話の録音も聞き取れるようになり、単語もどんどん増えていきます。英語やドイツ語のような、他人が書いた文法書や辞書からでなく自分で直接話者の人に聞くことでその言語のことを理解していくわくわく感に見事にはまってしまい、それから年に8回ほど調査に出かけるようになりました。当時宮古高校の校長だった仲間博之先生や公民館長だった仲間忠さんをはじめとする土地の人たちとも仲良くなって、京都大学にもなんども来てもらい、フィールドワークの授業にも先生として参加してもらいました。現在は仲間先生と岩崎さんとで池間方言の辞書を作っています。金田一の随筆を読んだから50年以上たってフィールドワーク調査を始めるとはなにか運命的なものを感じざるを得ません。



デジタル博物館「ことばと文化」では宮古島・西原地区の言語と文化を映像や音声で紹介しています

<http://kikigengo.jp/nishihara/>



フィールドワークの記録を記した『琉球列島の言語と文化』。宮古島西原地区の創作童話や、豊年祭の様子を収録したDVD付(2013年)

プロジェクトリーダーが語る 日本語の個性 ②

小磯花絵

KOISO Hanae

こいそ はなえ ● 音声言語研究領域 准教授。専門は認知科学、言語学、日本語学。国立国語研究所理論・構造研究系准教授などを経て、2016年4月から現職。

プラシャント・パルデシ

Prashant PARDESHI

理論・対照研究領域教授・研究情報発信センター長。専門は言語学・英語学。国立国語研究所言語対照研究系教授などを経て2016年4月から現職。

小木曾智信

OGISO Toshinobu

おぎそ としのぶ ● 言語変化研究領域教授。専門は日本語学、自然言語処理。明海大学講師、国立国語研究所准教授などを経て、2017年4月から現職。

— 先生方に共通する研究として、「コーパスの開発」がありますが、そもそもコーパスとはなんですか。

小木曾 簡単に言えば、実際に使われた言葉のデータをたくさん集めて、コンピュータで検索したりできるように蓄えて、かつ、特に言語研究のために必要な情報を付けたデータというふうに捉えておけばいいのかなと思います。要するに、たくさん実際に使われた言葉が入っているデータベースです。

— 開発しているコーパスにはそれぞれ違う部分があると思います。担当しているプロジェクトの内容も兼ねて教えてください。

小木曾 私は『日本語歴史コーパス』という名前の「通時コーパス」と呼ばれるものを作っています。古い時代から今までずっと時間を通して、時間の流れに沿って言葉を調べられるコーパスで、これを使えば、日本語の歴史が分かるといえるようなものを作りたいと思っています。万葉集の時代、奈良時代以前から大体近代、明治、大正ぐらいまでを対象として今コーパス化しています。

日本語の歴史は千年以上たどることができるので、千数百年分のいろんな時代のいろんな種類のテキストをコーパスにする、そういう仕事です。プロジェクトとしては、それを作るだけではなく、作ったものを使って、今までできなかったような研究

を行うことも目的としています。

小磯 私のほうは、日常生活の中で私たちがどういう言葉を使っているのかという、話し言葉を調べるプロジェクトです。家族と何かしゃべったり、仕事の場面で使う言葉であったり、同僚と昼食を食べたり、あるいは飲みに行ったり、友だちと会ったりとか、いろんな場面の日常の中で交わされる会話を収録します。

書き言葉と違うのは、収録をし、それを文字に書き起こして、整理をして公開をしていくという点です。一般の方に調査協力をお願いして、収録機材を2、3カ月所有してもらい、一人の人にいろんな場面の会話を録ってもらっています。それを40人

の方にお願いして、全体で200時間の会話のコーパスを作ろうとしています。例えば、自分のお母さんとしゃべるときと、旦那さんのお母さんとしゃべるときとは話し方が違う人もいます。そういうふうには、場面や相手によって、どういうふうには話し方が異なってくるかというようなことを研究することができます。

それ以外にも、50~60年前に、国語研が中心となって、井戸端会議やお魚屋さんなど、いろいろなところに行き行って録った話し言葉の音声データがあって、それと比べながら、50年ぐらいのスパンでどういうふうには話し言葉が変わってきたのかということも見ようとしています。

ブラシャント 「統語・意味コーパス」プロジェクトでは、新聞、聖書、教科書など種類はさまざまですが、現代語の書き言葉を対象にしています。統語的な構造（注：8ページの図をごらんください）を付けていて、言語の構造の研究ができるようなコーパスを作ろうとしています。

アメリカのホワイトハウスで報道官が、記者の歩きスマホを注意したんです。その記者はずっとポケモンをやっていて真面目に話を聞いてない。最後に皮肉ったジョークで、「Did you get one? (あなたは一匹捕まえたのか?)」と言うんです。その日本語の字幕は「捕まったの?」。動詞しか出てきてないんですね。「あなたが」「ポケモンを」という二つの項は隠れているけれど、潜在的にはそこにあります。

日本語のデータを英語で翻訳するときには、このような省略の問題が出てきます。だから省略されているものがわかるように、構造のついたコーパスを作ろうとしています。

それから、普通は、統語論の研究はほとんど内省で、都合のいい例文

を作っているわけですが、そうではなくて、実際に使われている言語に基づいて統語の研究もやるべきだというふうに思っています。

— 先生方のコーパスが、言語の研究、あるいは、一般社会において、どのように役立つかについて、ぜひ教えてください。

小木曾 われわれは日本語の研究者なので、まずは言葉を研究するためにコーパスを作っているのですが、特に歴史コーパスということになってきますと、過去千年以上伝えられてきた貴重な資料を、「今の技術で使いやすい形にしていく」ということをしていることになると思うんですね。

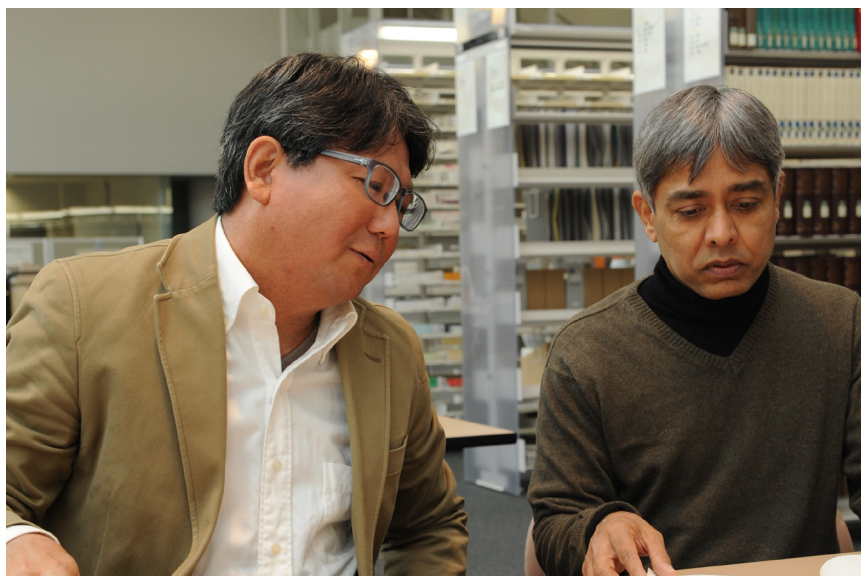
特にもう千年以上前の言葉ともなれば、われわれ日本人、日本語をしゃべるものたちにとってすごい財産であり、宝物だと思うんです。ただし、その財産や宝物を箱に入れてとっておいてもその価値は気付かれなくなってしまいうわけで、その時代、その時代で使えるように光を当てていくということは、大事なことだと思います。

今の時代、紙の本で読みましょうというだけではもうなくなっていきます。コンピュータを使って自由自在

に調べられるようなものを作っていくというのは、ただ言語研究者だけではなくて、日本語を話すわれわれみんなにとって、おそらく価値のあるものになるんじゃないでしょうか。一種の文化財というか、それをただ保存しておくというのではなくて、使えるようにしていくという、そういう価値があるのではないかと。

そういう社会にとっての価値という意味でいうと、国語教育の中では中学生や高校生は古文を習いますが、そういう授業で、今までのやり方とはちょっと違う、情報教材として古典教育などに生かすということも、このコーパスの役割として考えられて、実際にそんな研究もプロジェクトの中で進めています。まずは、中学・高校の先生方にコーパスを利用してもらうことから始めたいと思います。

小磯 今回のデータは、映像も一緒に集めているので、私たちがどういう生活をしていて、その中でどういう言葉を使っているのかという、その文化を伝えていくことにつながります。何十年かたつと、やっぱり生活も変わり、それで使っている言葉も変わっていく。そういう記録をずっと残していくというのは、私た





ちがどういう文化の中で育っているのかということを伝える、貴重な資料になるだろうなと思います。

教育という面で言うと、生活の中で、どういう文脈の中でどういう言葉を使うのかということが非常に重要になってきます。例えば、日本語を学習する人の多くが、「ドラマを見るとすごく勉強になる」と言うのは、いろんなシチュエーションで、どういう言葉を使っているかというのが分かるからだと思います。また、最近AI（人工知能）スピーカーとか、介護ロボットとか、家庭の中にどんどんAIのものが入ってきています。それらと会話をしていくことが必要になったときに、技術をきちんと確立していくためには、家庭の場面でどういう話し言葉を使っているのかということが、とても重要になってくるんですね。

ブラシャント 日本語は他の言語とどう違うんだろうか、どこが似ているんだろうかということを調べるのに役に立ちます。非常に簡単な日本語で、例えば、「太るお菓子」と言ったら、お菓子が太るわけではないということはみんな分かってい

るんですね。この文は、英語で翻訳するときに大変苦勞すると思います。だけど、「おいしいお菓子」というのは簡単に翻訳できるし、「昨日買ったお菓子」は簡単に翻訳できる。だけど、「太るお菓子」は難しいねと。同じお菓子が右側に来て、前に修飾語が付いていて、だけど構造が違っているということは、どこか教育の過程で学ぶ必要がある。そういうときに私のコーパスは役に立つのかなと思います。

— データを集めて、それに情報を与えていくという作業は、とても大変そうですが。

小磯 日々大変です。

ブラシャント 完成したコーパスしか見えていないのですが、その裏にはものすごく苦勞する場面があります。大学院生などがアノテーション作業を一生懸命に行っています。データは多種多様で、そのときそのときにいろいろ判断しないといけない。

小木曾 本には書いた人などの権利がありますので、その内容をコーパスにするときには、許諾をいただいて使うということが必要になるのが大変です。

それから、特に江戸時代から明治のものについては、当時出版されたままの資料をコーパスにしています。明治のころの本は、活字であっても今のわれわれには非常に読みづらいというようなこともあります。江戸時代ですと、いわゆる続け字の変体仮名を含むくずし字で書かれていますので、まず解読して普通の文字に起こしてデータにしていくという作業になります。そうやって普通に読めるテキストのデータができた後に、それぞれに単語の情報を付けるとこ

ろまでやっています。

全部のテキストに単語の切れ目を入れて、それぞれがどんな品詞でどんな読みでという、形態論情報と呼ぶものを付けます。それには形態素解析というコンピュータを使った言語処理の技術を使うのですが、古い時代のテキスト用の辞書は存在しないので、自分たちで作って、技術開発も行ってなんとかしなければなりません。そして、機械が処理をしたデータに人で直すという処理を加えて、最終的になんとか使える形になったものを公開するというをしています。大変です。

ブラシャント 機械が7割正しいとき、3割は人間が直さないといけない。一つ一つセンテンスごとに見ていって修正して、また機械にそれを学習させる作業をやらないといけないので、やっぱりアノテーション現場で行っている作業は、非常に大変です。

小磯 音をとったものを聞いて、「何を言っているのか」を考えること自体が、解釈することなんです。みんながきれいに発音して明確な文章を発話していれば、そこには解釈は入らないかもしれませんが、普段の話し言葉では、音のレベルで言い間違えたり曖昧な話し方とかするので、それをどういうふうに書いていくのかということも本当に解釈になるので、まず文字にするというところからして、実は研究なんですね。

それに加えて、単語に区切って品詞を付ける、そこまでは書き言葉のコーパスでもするんですが、音声の研究をやろうとすると、イントネーションも必要です。ここは上昇調のイントネーションだよとか。あるいは、「なんかでー」「なんかするどー」というように若者は上昇下降

の音調を使うと言われますが、実際には、年配の方もかなり使うので、そういう音調を付けたりとか。

また、一部のデータには、今のフレーズが質問で、これはそれに対する応答であるとか、一部のデータには、そういう「行為」の情報を付けます。また、文節間の係り受けの関係などいろんなアノテーションが考えられます。どれも一個一個非常に大変です。ただ、そういうふうにいるいろんなアノテーションがあると、例えば統語と韻律の関係をみるなど、様々な観点から研究することができるので、一部のデータには、いろんな情報を付けてます。

小木曾 ちなみに、「アノテーション」という言葉は、われわれはよく使います。本文となるデータの上にあるいろんな情報を重ねて付けていくというものです。単語の情報もあれば、統語情報、文法的な情報もあれば、いろんなイントネーション、韻律の情報とか、そういうものを重ねて付けていきます。そういう「アノテーション」を付けたものをコーパスとして公開していくことになります。

— 「アノテーション」の難しさをもう少し聞かせてください。

ブラシャント 言語には様々な現象があります。それを全てアノテーションしているわけではありません。それをやると、たぶんわれわれの人生では終わらない。たぶん次の人の人生でも終わらないと思う。アノテーション作業は客観的かつ一貫性を保ちながら行わないといけない作業であり、かなり時間かかるということをみんなに理解していただきたいなというふうに思います。

小木曾 コーパスに対するアノテ

ションというのは、大学院ぐらいまで行って日本語学、言語学の研究してきた人でないとできないことが多い。

ブラシャント 人同士で意見が違ったりすることもあります。そのときには、みんなで議論をして一貫性を保ちながら客観的にやる。一貫性をもってやるというのが非常に大変な作業です。

小磯 アノテーションをするために基準を作るんですよ。その基準を作るというのは、ほとんど研究なんです。実は研究なので、結構みんな苦労してる。けど、ある意味楽しいところでも、もちろんあります。

— コーパスを作って、それを使って研究するだけでなく、作る過程で研究が生まれているんですね。

聞き手 ● 折田知之さん（東京外国語大学大学院生）



日本語の構造を探る

プラシャント・パルデシ

Prashant PARDESHI

●教授／専門は言語学、言語類型論、対照言語学。神戸大学大学院文化科学研究科博士課程修了、博士(学術)。2009年に本研究所着任。国立国語研究所第1回所長賞、第12回特別所長賞受賞。



言葉の規則

私たち人間が日常物事を考え、互いにコミュニケーションを行うにあたって、言葉は欠くことのできない手段です。人間は幼児期から周囲で話される言葉を自然に覚え、それがその人の母語となります。これに対して、日本人が英語などの外国語を習い、また外国人が日本語を学ぶときには、「文法」と呼ばれる言葉の規則とともに学習する必要があります。

言葉を自然に習得し使用するときには文法は意識されず、見えないところで働いているのですが、外国語として学ぶ際には表面に出て来ることになります。現代ではコンピュータによる言語情報処理が実用化されようとしています。そのように機械に言葉を「教える」ためにも文法が必要です。

文の構造とコーパス

私たちは学校で主語や述語、また関係代名詞を伴う文による名詞の修飾などを習いますが、これらの文法知識に共通していて、それらを支えているのが、文(センテンス)を構成している単語やその集まりである句が互いに修飾したり修飾されたりする関係を持っている、という事実です。このことは、「文は構造を持つ」と言い換えることができます。文の構造に関する規則が文法に他なりません。このように構造を持つ文を話してコミュニケーションを行うことは、質的にも量的にも限界のある脳を使って周囲のありとあらゆる事柄を表現するために人間が進化の過程で身に付けた能力です。

例えば、「山の緑が美しく見える。」という文は、おおよそ、

[[[[山] の] 緑] が] [[美しく] 見える]

のように緑のカッコのまとまりごとに構造をなしています。このうち、一番外側のボールド体で示すかっこの部分、「山の緑が」と「美しく見える」は、それぞれ主語と述語に相当します。このような構造付けされた言語データを集積することで、簡単に言えば「文は主語と述語からなる」というような文法規則を抽出することができます。また逆に、そのような規則をコンピュータに与えて、主語と述語からなる文を作らせることもできるようになります。

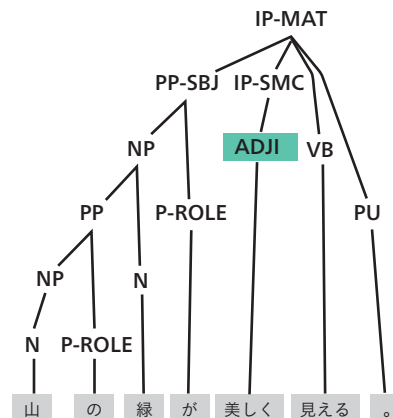
元々人工知能の研究によって生まれた概念に「句構造」というものがあり、近年になってこれを文構造の表示に利用した様々な研究が行われてきました。句構造を使えば文の意味が正しく扱われることが知られています。

新聞記事や小説などのテキストに文法情報を注釈として付加した言語データベースはコーパスと呼ばれ、多様な言語研究に利用されています。句構造をテキスト中の文に付加したコーパスは、文法規則や文型が現実にもどのように使われているかを解明するための強力な手段であり、英語をはじめとする世界の主要言語について整備が進められています。しかし、日本語についてはそのようなコーパスはこれまでに存在しませんでした。

私たちのプロジェクトで開発しているコーパスは、文の句構造に関する情報を持つ初の日本語コーパスであり、意味表示機能も兼ね備えています。こ

れによって日本語の文法研究のための革新的なツールを提供することができます。また、外国人に対する日本語教育やコンピュータによる日本語情報処理にも貢献することができます。

私たちのコーパスで提供されている、上で引用した文の文法解析情報を「木構造」の形式で掲載しておきます。



コーパスの構築と利用

このようなコーパスの開発のためには、日本語文法の研究者だけでなく、コーパス言語学や言語情報処理など、様々な分野の研究者が知見を持ち寄って共同研究を行う必要があります。そのため国内の大学等の研究者と共同研究体制を作ると同時に、海外の研究機関との連携・協力を進めています。また、日本語文法の研究だけでなく、外国人に対する日本語教育や言語情報処理へのコーパスの応用も今後積極的に進めていきます。すでにできている部分については、以下のサイトで公開しているので、ぜひ試してみてください。
<http://npcmj.ninjal.ac.jp/>

コーパスで知る日本語史

山崎 誠

YAMAZAKI Makoto

やまざき まこと ●教授/専門は計量日本語学。東京学芸大学大学院修了、博士(学術)。1984年に本研究所在着任。著書に『テキストにおける語彙的結束性の計量的研究』。言語処理学会第12回大会優秀発表賞受賞。



コーパスに触れてみよう

日本語研究に本格的にコーパスが使われるようになって十数年ほどたちました。若い研究者を中心にコーパスの利用が広がってきており、コーパスを利用した研究も数多く公開されています。しかしまだ、コーパスはハードルが高いと感じている方もいるようです。そのような場合は、まず自分の興味のあることを確かめるような使い方から入ることをお勧めします。とくに、日本語史はもともと文献資料に基づく研究分野ですから、コーパスとの相性は非常によいと言えます。また、検索用のツール「中納言」を使えば、プログラミング等の必要がないため、比較的直感的に知りたい結果を得ることができます。



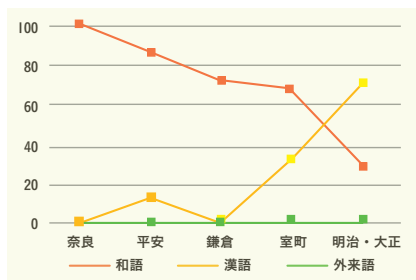
「中納言」(<http://chunagon.ninjal.ac.jp>)

いつごろから漢語が増えたか

日本語の語彙には、大きく分けて、和語、漢語、外来語という3つの語のグループがあります。これらはそれぞれ、大和言葉、(古代)中国語、西洋語(主として英語)、に対応しており、表記の面からも容易に区別できるものが多いです。このような語の違いのことを語種ごしゆと言います。日本語研究の多くの場面で利用されています。漢語は、もともとは「外来語」なのですが、取り入

れられてからの年月が長く、語数も多いことから、和語と並ぶ一つの大きなグループを構成しています。その漢語がどのようにして日本語語彙の中で勢力を伸ばしていったか、コーパスを使って調べてみましょう。

下の図は、「日本語歴史コーパス」を使って、時代別に名詞における語種の割合の変化を示したものです。割合は延べ語数で算出しています。



名詞における語種の割合の変化

「日本語歴史コーパス」には、2018年1月時点で奈良時代から明治・大正時代にいたる34作品、約1456万語(句読点などを除く)が収められています。図1からは室町期(資料としては「虎明本狂言集」)から漢語が増え始め、それと反比例して和語が少なくなっていることが分かります。なお、外来語は、明治・大正時代には約1.1%ですが、現代では、約8.8%に増えています(「現代日本語書き言葉均衡コーパス」による)。

古典におけるコピュラ文

次に、平安時代の随筆「枕草子」の有名な一節「春はあけぼの」を取り上げます。現代語でも容易に意味が通じるこの文は、「AはBだ。」という単純な構造であり、コピュラ文と呼ばれます。コピュラ文にもいくつかのタイプがありますが、上記のようなコピュラ

文がいつごろからどれくらい使われていたのかを中納言で調べてみます。以下の表に結果を示します。これは、中納言で、「名詞十は十名詞十。」と「名詞十は十名詞十なり十。」という2つの検索を行い、その結果を足し合わせたものに基づいた表です。多くの古典の本文には句読点は存在しませんが、「日本語歴史コーパス」では、『新編日本古典文学全集』(小学館)などの校訂本文に拠って句読点を施しているため、この情報を利用します。

表 コピュラ文の数

時代	作品名	ジャンル	用例数
平安	落窪物語	作り物語	1
平安	枕草子	随筆	12
平安	源氏物語	作り物語	4
平安	紫式部日記	日記	3
平安	大鏡	歴史物語	1
平安	讃岐典侍日記	日記	1
鎌倉	今昔物語集	説話	8
鎌倉	宇治拾遺物語	説話	3
鎌倉	十訓抄	説話	1
鎌倉	徒然草	随筆	3
鎌倉	とほすがたり	日記	2
明治・大正	国民之友	非文芸	7
明治・大正	女学雑誌	文芸・非文芸	16
明治・大正	太陽	文芸・非文芸	165
明治・大正	女学世界	非文芸	6
明治・大正	婦人倶楽部	非文芸	1
	合計		234

表からコピュラ文が意外に少ないという印象を受けます。平安時代の作品では「枕草子」が12例といちばん多くなっています。「枕草子」の7倍近い言語量がある「源氏物語」では4例と少ないこと、また、「源氏物語」と同じ作者による、「紫式部日記」では3例使われていることを考え合わせると、このタイプのコピュラ文は作り物語よりも随筆や日記になじむ文体的な特徴があったのではないかと想像されます。

気になったことをすぐに調べられるのもコーパスの利点です。試しにいろいろな検索を試みてください。

会話から日本語を見る

小磯花絵

KOISO Hanae

こいそ はなえ ● 准教授 / 専門はコーパス言語学・コミュニケーション科学。奈良先端科学技術大学院大学博士後期過程修了、博士（理学）。1998年に本研究所着任。



子供のころ、家に電話がかかってくると、たいてい母が出ました。少し話し出すと、その口調から、どのような関係の人が電話をかけてきたか、予想がついたものです。母の場合、自分の妹とはかなりくだけた言葉づかいで声はずんだ感じになりますが、姉だと親しさの中に丁寧な言葉が混じります。父方の親戚だとかなり丁寧になります。このように私たちは、話す相手によって、言葉づかいを使い分けることがあります。こうした使い分けは、会話の相手だけでなく、場面や話題なども関係してきます。

「俺」「僕」「私」の使い分け

例えば、自分自身を呼ぶ際に用いる自称詞。男性の場合は「俺」「僕」「私」などがよく使われますが、同じ人がいつも同じ自称詞を用いているわけではありません。

ある30代男性の事例を見てみましょう。ご自身に様々な場面での会話を収録してもらい、それを調べてみると、妻や仕事の後輩には「俺」を、妻の母や妻の友人には「僕」を、仕事の先輩や年上の知人には「私」を使っていることが分かります。「俺」や「僕」は私的な関係、「私」は公的な、改まった関係の相手に使っているようです。また同じ相手であっても、場面による使い分けが見られます。妻と二人で話す時は「俺」を用いるのに対し、妻の母がいる場で妻に話す時には「僕」を用いています。「俺」の方が「僕」よりもぞんざいな印象を与えるため、妻の母のいる場では妻に対しても「僕」を用いているのでしょう。

60代男性の場合も同様に複数の自称詞を使い分けています。自分の弟や妹には「俺」を、娘婿を含む家族との会話では一貫して「私」を使っています。地域ボランティアの会の役員会という、親しい仲間同士とはいえ改まった場面では「僕」を中心に使っていますが、2例だけ「俺」を使っている箇所がありました。ボランティア主催のイベントにやっかいな人が参加して腹が立ったという語りをしている箇所です。ぞんざいな印象を与える「俺」を用い、腹の立った経験の語りを演出していると考えられます。同じ相手、同じ場面でも、話題によっては異なる自称詞を使うことがあるようです。

日常会話コーパスとは？

こうした場面や相手による言葉の使

い分けを細かく調査するには、多様な日常場面の会話を収めた「ことばのデータベース」（コーパスと呼びます）が有効です。

「日常会話コーパス」プロジェクトでは、一般の協力者にご自身の日常生活の中で生じる会話を収録してもらい、コーパスとして整備しています。下の図にあるように、自宅での家族との会話や職場での会話、友人との会話、理髪店での会話など、多様な場面・相手との会話が集まってきました。先ほどご紹介した二人の男性の自称詞の使用は、このコーパスを使って調べたものです。来年度、全体目標200時間の会話コーパスのうち、先行して50時間の会話を公開する予定です。ぜひ、私たちが普段どのような場面でのどのような言葉づかいをしているのか、調べてみてください。



『日本語日常会話コーパス』に納められている会話の例

ことばの多様性と コミュニケーション

そうじゃ そうですわよ そうだよ そや そうよ んだ ソウアルヨ



平成30年2月3日、東京証券会館にて第12回NINJALフォーラムが開催されました。

開催日直前に降雪に見舞われたにも関わらず、360名を越える事前申込みがあり、盛況の一日でした。

このフォーラムは、国語研究所が主体となって実施する研究や、他機関との連携による研究成果を、一般の方々にも知っていただくことを目標に、毎年度開催しています。

ことばの多様性の再発見

地理的バリエーションでは、沖縄各地の言葉などの日本の消滅危機言語を、次世代にどう継承していくか、どのように方言が生まれ、広まっていくのかということについて解説されました。

社会的・機能的バリエーションでは、「千と千尋の神隠し」の登場人物の言葉遣いから、そのキャラクターが物語の中でどのような存在であるかという分析が示されました。また、テレビで馴染み深い「家庭教師のトライ」のCMを例に、言葉と「状況によって変わる人物タイプ」であるキャラクターとの関わり方が多様であるということが説明されました。

外国語との係わりにおけるバリエーションでは、日本語学習者には発音や文法に母語の影響がみられる「お国柄」があることや、相手を気遣うていねいさを表したいという気持ちは世界共通であっても、その方法は多様性に満ちていることが紹介されました。

「どうなる？」から「どうする？」へ

司会の「これからの日本語はどうなるか」という問いかけに対し、「次世代へ方言を継承していくためには、若者の言葉遣いに口うるさく文句を言ってはいけない。多少違っていても、方言を話そうとする姿勢に耳を傾けるべきである。」(田窪)、「言葉は単なるツールではなく、文化そのもの。グローバル化に伴い道具としての言葉のイメージが強まる中、文化としての言葉の役割を強調していきたい。」(定延)などの発言に、来場者も深くうなずいている様子でした。

た。パネルディスカッションの後半は、来場者との活発な質疑応答も行われました。

日本語の多様性とコミュニケーションについて 深く考える機会に

終了後は来場者から「もっと言葉を意識して、言葉を自ら選んで発信したい」「今まで、言葉はツールという認識が強かったが、それだけではダメなのかもしれないと思い直す機会になった」「言葉の持つ差異性と共通性、相反する性質によって多様さが生まれる。そのおもしろさに触れることができた」などの声が聞かれました。

一口に日本語と言っても、方言や年齢、性別といった社会的なバラエティ、そして外国人が使う日本語など、実に様々な変種があります。それらはどこから、どのようにして生じるのか。そのような多様性は、私たちのコミュニケーションにどのような影響を及ぼすのか、日本語の多様性とコミュニケーションについて来場者と一緒に深く考える機会となりました。

プログラム

第1部：リレー講演

Part 1 地理的バリエーション

- 講演①「方言はどこまで通じるか？」(田窪行則)
- 講演②「方言の生まれるところ」(大西拓一郎)

Part 2 社会的・機能的バリエーション

- 講演③「ポップカルチャーと役割語」(金水敏)
- 講演④「ことばとキャラ」(定延利之)

Part 3 外国語との係わりにおけるバリエーション

- 講演⑤「日本語学習者のお国柄」(石黒圭)
- 講演⑥「ていねいさは世界共通か？」(宇佐美まゆみ)

第2部：パネルディスカッション

テーマ：「どうなる？これからの日本語」

コーディネーター：野田尚史

パネリスト：田窪、大西、金水、定延、石黒、宇佐美





研究者紹介 007

前川喜久雄

音声言語研究領域 教授

コーパス整備の転換点に ちょうど私がいたんです

まえかわ きくお ● 1956年京都府出身。鳥取大学講師などを経て、国立国語研究所には1989年から在籍。『講座日本語コーパス』（朝倉書店全8巻）の監修を務める。2011年、2012年日本音声学会 優秀論文賞を受賞。

— 音声学に長年関わっていらっしゃる印象があります。

学部の卒論から40年近く、音声学をやっています。400本近い論文を書きましたが、その9割弱は音声に絡んでいます。また、1999年に言語資源の設計や構築の仕事为国語研の仕事として任されました。言語資源とは、CSJ（日本語話し言葉コーパス、東京工業大学・情報通信研究機構との共同研究）やBCCWJ（現代日本語書き言葉均衡コーパス）と呼ばれるものですが、コーパスの構築を2016年まで続けました。

— 今一番興味をもっているご研究は？

言語資源、つまり、バランスの取れた大量のデータであるコーパスを使った自発音声の研究をするというのが、一番の興味の対象です。

音声研究は実験をかなり重視します。ただ、実験は仮説があって初めて成り立ちます。ということは、仮説に関する部分だけを見るわけです。つまり、実験というのは、狙っている対象をコントロールしなければなりません。

ところが、話し言葉の中には、コントロールできない大事なコミュニケーション情報が山ほどあります。子音とか母音だったら、文字に書いて読んでくださいといえれば大体そのとおりになりますが、イントネーション、口調や感情はそういうわけにはいきません。また、言い間違いを研究しようとして、「こう言い間違えて」と言うと、それは言い間違いではないわけです。

それらを研究するには、話しているデータを集めることが必要で、ものすごくデータの散らばりが大きいから、それを大量に集めなければいけないんで

す。大規模なデータがいったんできれば、実験では見つからない部分を、信頼性を持って分析することが可能になります。

— 自発音声の音声を解明するためにコーパスを作ったのでしょうか。

いや、実験を重視した研究に疑問を持ちはじめたのはアメリカから帰った後の1995年の頃です。この頃からパラ言語情報（イントネーションやリズムなど）を研究しはじめて、そのあとに偶然にも、99年からCSJをつくることになりました。CSJはももとは音声認識のためのコーパスでしたが、言語学の研究にも使えるように、いろいろ情報を付与して作ったんです。

あまり知られていませんが、国語研究所は、世界的にも古く1950年代からきちんとサンプリングされたデータを作っていました。ただそれは、今流に言う言語資源ではありません。作っただけではなく公開し誰でも使えるようにならなければ言語資源とはいえません。その第1号がいわばCSJで、その後関わったBCCWJもそうですが、そのあとは大体作ったものを公開する流れになってますよね。

— 言語系の学会では、コーパスの研究が多くなりました。流れの大本に、先生が関わっていたんですね。

誰かがやらなきゃいけないので。ただ、CSJとBCCWJは、少し性格が違うんです。BCCWJは、外国にはきちんとしたコーパスがあるんだけど、日本にないから作らなければ、という考えで進めました。CSJはそれとは違います。世界で誰もやったことのないことをしようと思って作りました。

今は、皆さんが使っているスマホを通じてGoogleなどが桁違いの音声データを集めていて、音声認識用のためにデータを集めるという考えはなくなりました。例えば、スマホで認識しようとして間違っていたら、正解が出るまで話し続けますよね。どれが正解かをシステムに伝えているわけです。

ただ、それは音声認識のためのデータであって、話し言葉についてきちんとした付加情報、つまりアノテーションが付いているものはなく、CSJがいまだに世界で一番大きいんです。情報検索に使えるアノテーションの質が高く量が圧倒的に多いんです。CSJのユーザは工学系の人たちで、彼らはいくらでも好きなように利用しています。

BCCWJは、機械翻訳の人も使いますが文系の人が多いようです。5～600本は論文が出てると思いますが、分析が物足りない感じがします。その辺りは学会の課題だろうと思います。

言語資源がある程度できてきましたし、統計技術の理論も非常に進歩してきたので、ようやく言語学の面、統計学の面の双方からみて、いい時期になっている感じがします。次の世代に、新しいブレイクスルーをつくる人が出てくるんじゃないかという気がします。

そのためにこの研究所が役目を果たせるといいと思います。天才を生むことはできないけど、言語資源をつくる、整備することは地道にやってけばいいわけです。大学ではなく国語研だからこそできることがあると思います。



研究者紹介 008

ビャーケ・フレレスビグ

共同研究員・オックスフォード大学 教授

古い日本語に ネイティブスピーカーはいない 皆、同じ立ち位置で

Bjarke Max Frellesvig ●1961年デンマーク出身。コペンハーゲン大学、オスロ大学などを経て、1999年からオックスフォード大学教授。現在、同大学日本語研究センター長。「オックスフォード上代日本語コーパス」を構築。また、ヨーロッパ日本研究協会の会長も務めた。

— デンマークご出身の先生がなぜ、 上代日本語（奈良時代以前の日本語） にご関心？

デンマークのコペンハーゲン大学で、言語学のコースをとりました。大学では、ヨーロッパ語ではない語を習わなければならない、高校を卒業後に日本に行ったこともあり日本語を選びました。21歳のときです。

僕は、歴史言語学に興味があって、日本語だけでなくその歴史も知りたかったと思っていました。大学では、古文の授業もあり、そこで竹取物語や伊勢物語に出会ったことが、上代日本語に関心をもつきっかけになったと思います。

— 現在、コーパスを作っている しゃるとのことですが。

奈良時代の日本語のコーパスです。あるテキストと、テキストの情報を一緒にまとめ、コーパスにしています。

2008年から私が勤めるオックスフォード大学（英国）で各国の古い言語のコーパスを作るプロジェクトがはじまりました。2010年から国語研と共同研究をするようになりました。コーパスは国語研でもうすぐ公開する予定です。

コーパスにしておけば、調べたいことが簡単に調べられます。上代の日本語は全部漢字で書かれたものです。一つ一つの漢字がどうやって使われているかは、電子テキストがあれば、簡単に調べることができるんです。紙のままではできませんよね。

私は、格助詞「が」と「の」に関することはとても面白いと思っています。現代語では「が」は主格、「の」は属格（所有格）ですね。ですが、奈良時代の

日本語では同じ所有格の格助詞でした。それらがどういふふうに使分けられていたのかについて、コーパスがあれば、簡単に調べることができるんです。

— 国語研にも万葉集のコーパスが ありますよね。

オックスフォードのコーパスには、シンタクス（文の構造）の情報も入っていてその点で大きく違います。コーパスから自動的にツリー（図参照）を出すことができます。ローマ字表記が少し見にくいのですが、慣れたらすごく便利です。たとえば、和歌の構造などをすぐに見ることができるからです。

— 日本語、しかも古い日本語が海外 で研究されているのに驚きます。

それはよく言われます。でも、よく考えたら、現代日本語であれば、母語話者は自分の直観とかを使うことができますが、外国人は無理ですよ。でも上代日本語の場合は、母語話者、ネイティブスピーカーはいないので、皆が同じ資料を見て、いろんな角度からその資料やテキストを解釈することができるんです。

— 今後のご自身のご研究について 教えてください。

「主語はどういふふうに表示されているか」について関心をもっています。現代日本語では、格助詞とか係助詞を使いますが、上代日本語では、なんの助詞も無い場合もよくあ

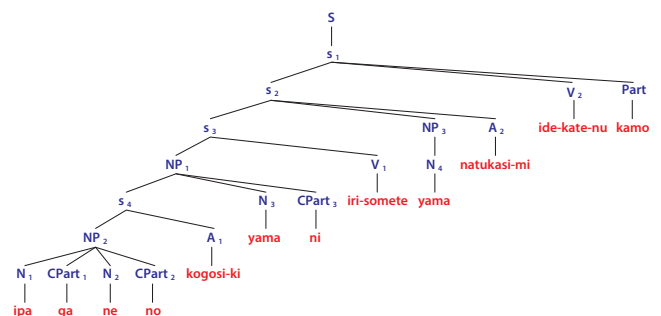
ります。その文法や、どういう場合にどの格助詞を付けるかなどについては、基本的な研究課題ですが、まだ完璧に分からないところがいっぱいあると思います。

語彙の構造にも興味があります。上代日本語のテキストより前の記録はありません。僕はいわゆる原始日本語の再現にも興味があるんです。コーパスには辞書を取り出す機能があるので、その辞書を使えば、その語彙の構造の研究に役立つのではと考えています。

簡単な例を言えば、現代日本語の「はじめる」や「はじまる」は、その単語が奈良時代にもありましたが、それより「そめる」という単語を使っていたようです。今も残っていると思いますが。

— 「見そめる」とかの「そめる」？

そうです。でも、その「そめる」は、実は「はじめる」の基礎だと考えています。「そもそも」という副詞はその「そめる」と関係があると考えられます。一番古い記録の「そめる」「そもそも」「はじめる」からたどっていけば、そういう関係が見やすくなると思います。



万葉集7・1332

読み：「岩が根の こそしき山に 入りそめて 山なつかしみ 出でかてぬかも」を分析したもの。ローマ字で表現されている。たとえば、赤字の「岩 ipa」の部分がN(名詞)であり、「岩が根の ipa ga ne no」がNP(名詞句)であることがわかる。



研究者紹介 009

蒙 韞

日本語教育研究領域 プロジェクトPDフェロー

外国人が日系企業で働くときの言葉の姿を捉えたい

MENG Yun ●中国・広西チワン族自治区出身。岐阜大学・名古屋大学で日本語、日本語教育を学ぶ。2010年3月に博士号取得。2016年4月から現職。

— どういったご研究を？

外国人が日本語を勉強するとき、どんな問題があるか。どうしたらうまく上達できるかということを研究しています。大学院の時代は、外国人向けの日本語教育を専門として勉強しました。今の仕事もそれと関わっていて、とても面白いなと思っています。

以前は「聞く」「話す」を中心に研究していましたが、今は、「読む」「書く」も見ています。

— それは文法的なことですか？

文法というよりは、談話分析と会話分析のほうです。名古屋大学には当時大曾美恵子先生や尾崎明人先生をはじめ、談話分析の分野で有名な先生がいました。おそらく他の大学に比べると、これらの分野の授業が多かったのではないのでしょうか。授業内容が面白かったこともあり、自分の修士と博士のときのテーマとして研究してきました。

また、社会言語学とも少し関わりますが、アカデミックというよりは、実用的なことに焦点を当てて考えたいと思っています。例えば、外国人の会社員が日本・海外の日系企業で仕事をするときの言語のコミュニケーションの問題に関心があります。

— 自分の研究でここが面白いというところは。

ポライトネス（配慮表現）や語用論的（表現の捉え方）な問題です。とても日本語が上手な人なのに、ある場面になると何か変になるケースがありますよね。例えば、遅刻する場合は、日本人はまず謝ります。でも、外国人は理由からだらだらと説明するので、日本の方から見ると、「この人、本当に悪

いと思っているのかな」と思うわけです。こういう点はなかなか学びにくいです。

自分も実は、学生時代に会社で研修生の翻訳・通訳のアルバイトをしていて、同じような説明をして日本人の上司によく怒られていました。

— 大変だったことがたくさんありそうですね。

実は山ほどあります。例えば、ディスカッションをするときには、中国人は、自分の意見をストレートに出します。相手のことにすごく反論して怒られても、何か結論が出てくればいいと考えます。

日本人の場合は、相手との人間関係を維持するため、できれば相手と直接ぶつからないように相手に合わせたり、本当は正反対の意見を持っていても、今この場ではもう議論をやめて後でお話をしましょうとなったりします。

問題を解決するために自分の考えを率直に言っても「後でお話をしましょうか」と言われて、本当に後で話してくれるかなとすごく期待をしたのに、何もないようなことになっていて、「え？」っていう経験もあります。どうすればいいかなと（笑）。

— 今は、ビジネス日本語のほうにご興味があるのでしょうか。

二つの柱があります。一つは「アカデミック日本語」で、そちらのほうは「書く」と「読む」、つまり、文章理解の研究をやっています。例えば、一つの文章を読み、知らない単語がたくさん出てくるとき、何かを手掛かりにして推測すれば、知らなくてももううまく理解できます。その習得のメカニズムを

研究しています。

もう一つの柱が「ビジネス日本語」です。その中でさらに二つにわかれます。

一つは、研究所の石黒圭先生の下で「クラウドソーシングと表現」について研究を行っています。例えば、求人情報を書くときに、「この仕事内容についてどういうふうに行けば応募しに来る人にとって分かりやすいか」ということを取り扱います。

外国人がこのクラウドソーシングを利用するケースも増えてきています。発注は日本国内ではなく海外になる場合もあります。海外のワーカーに日本語の発注文書をどう分かりやすく読んでもらえるかということも考えています。

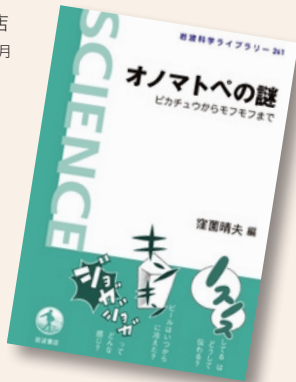
仕事が決まればその後、発注者と受注者の間でも何回もやり取りができてきます。やり取りの日本語も研究したほうがいいかなと思っています。効率よく、お互いに理解できれば、次の契約につながるのではないのでしょうか。

もう一つは、大学院での研究の続きになりますが、実際に日系企業で働いている外国人を対象に、自然会話のデータを取ることを考えています。

「仕事においてどういう問題があるのか」ということについて多くの外国人の方にインタビューをして、分析したいと思っています。

オノマトペの謎 ピカチュウからモフモフまで 窪園晴夫 編

岩波書店
2017年5月



本書は、2017年1月に国立国語研究所で開かれた第10回NINJALフォーラム「オノマトペの魅力と不思議」に基づいた報告書である。ここでは「事件」として紹介したい。

今回あらたに全国に指名手配されたのは、通称・小野窓平。容疑は共謀罪。音象徴という武器を集め、日本の言語体制に揺さぶりをかけようと謀議していたことが疑われたのである。事態を重く見た日本語政府は早速、国立国語研究所に捜査を命じた。陣頭指揮に立った副所長の窪園氏は、国内外から7名の言語プロファイリングのプロを招集し、犯人像の究明を託した。これによって、当初は幼稚・

単純な愉快犯と見られていたが、じつは驚くべき周到な知能犯であることが、さまざまな観点から浮き彫りにされた。何よりの証拠が音象徴を武器とする点である。それは言語に遍在するが、日本では小野窓平なる者の手にかかれれば、容易に誰でもどこでも利用可能であり、その破壊力は計り知れない。この武器がゲリラ的に集中使用された場合、既存の概念世界に安住する人々を大混乱に陥れるのは間違いない。それを未然に防ぐには、すぐにでも多くの捜査協力者が必要であると、本書は強く訴えている。それが叶うならば、小野窓平の謎の解明もそう遠くはあるまい。 ▶半沢幹一（共立女子大学）

現場に役立つ日本語教育研究③ わかりやすく書ける 作文シラバス 石黒圭 編

くろしお出版
2017年12月



作文指導ないし作文添削というのは、日本語教師には避けて通れないものである。一方で、人様の文章をどうこう言えるほど、自分は文章がわかっていると言えるのだろうかという悩みも併せ持つ。こうした不安に答えるため、本書では日本語母語話者、中国人日本語学習者、韓国人日本語学習者がそれぞれ書いた計180本の作文をデータとして分析を進める。本書の最大の特徴は、実際の作文（しかもジャンルをまたぐ横断的データ）に基づいているという点である。ただし、データを計量的に見るだけでは、良い作文とは何かは見えてこない。本書では「読み手に優しい文章」の条件として、

- ①正確で自然な日本語
- ②流れがスムーズな日本語
- ③説得力のある日本語

の三つを設定して議論が進められる。作文だけではなく、それを読む人間も見ようというわけである。しかも、そこに学習者の「今」だけではなく、初級、中級、上級という学習者の成長過程を縦断するシラバスをにらんでいる点が素晴らしい。

漠然と抱えている作文添削への不安を解消するため、見取り図を与えることにより、「実はこうなっていた」を見せる。本書はそんな論文集である。

▶森篤嗣（京都外国語大学）

A Grammar of Irabu: A Southern Ryukyuan Language Michinori Shimoji

九州大学
出版会
2017年2月



文法のあるべき姿を示すものとして、すべての言語研究者に本書を強く薦めたい～田窪所長も帯に書いているが、これにつきと思う。本書は著者が2008年にオーストラリア国立大学に提出した博士論文の加筆修正版であり、南琉球宮古伊良部島方言を独立した言語体系としてとらえ、その全体像を体系的・包括的に記述した参照文法である。国外の状況とは異なり日本ではそれまで若手研究者が参照文法を博士論文として執筆することはなかったが、著者の博士論文以降、多くの若手研究者によって琉球諸語の参照文法が博士論文として執筆された。詳細な目次を見てもわかるように

「参照文法とはどのようにして書くか」が明確に示されており、このような方法論を本書によって国内の研究者たちが手にしたと言えるだろう。また本書は言語理論を持ち、かつ曇りなき眼で一つの言語の記述に臨むことの重要性も伝える。例えば伊良部島方言の韻律は高低が繰り返しあらわれるという特徴を持つが、著者は日本のアクセント研究の諸前提を一度すべて捨てて再考することで正確な記述を可能とただけでなく、英語のストレス規則の一般化とも関連させ、韻律類型論におけるトーン言語とストレス言語の壁を打ち破りもしている。

▶山田真寛（国立国語研究所）

編集後記

国語研究所は2017年10月に新たに田窪行則所長が着任しました。今号では、新所長の自己紹介、コーパス構築に関連した3プロジェクトのリーダーの鼎談などを中心にお送りしました。

表紙の写真は前号に引き続き、「海」にまつわるものになりました。「えっ？」と思って見返した方もいらっしゃるかもしれませんが、これは国語辞典『言海』の1ページなのです。『言海』は大槻文彦がたいへんな苦心の末に独力で完成させた日本最初の近代的な国語辞典。その苦労は、この辞書の後書きにあたる「ことばのうみのおくがき」に述べられています（「青空文庫」にも入っているのでぜひ読んでみてください）。

写真は国語研究所蔵、1891（明治24）年刊の一冊本の『言海』から、「ことば」を載せたページを選びました。中段右から三つ目の項目が「ことば（名）言葉」ですが、「こ」の字が「古」の崩しからできた変体仮名になっています（現在使われる「こ」は「己」からできたもの）。

変体仮名は、今ではそば屋の看板や割り箸の袋以外ではあまり見かけませんが、江戸時代以前はもちろん、明治時代以降の活字でも使われていましたし、現在でも学術的な目的等で利用される場合があります。そこで、変体仮名を標準化してUnicodeに入れ、パソコンで使えるようにする取り組みが行われましたが、ここでは国語研究所の研究者が中心的な役割を果たしました。初期のJIS漢字コード以来、国語研究所はこのような標準化活動においても大きな役割を果たしています。

（小木曾智信）

次号予告

言語資源の整備と 研究成果発信

コーパス開発センター
研究情報発信センター

国語研 ことばの波止場 vol.3

平成30(2018)年3月30日発行

編集 国立国語研究所研究情報誌編集委員会

発行 国立国語研究所
〒190-8561
東京都立川市緑町10-2
電話042-540-4300(代表)

協力 くろしお出版

デザイン 黒岩二三[Fomalhaut]

無断転載を禁じます

©National Institute for Japanese Language and Linguistics