国立国語研究所学術情報リポジトリ

NINJAL Research Digest vol.10 (2021.9)

メタデータ	言語: jpn
	出版者:
	公開日: 2021-10-29
	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/10.15084/00003469

とばの波止場

NINJAL Research Digest



特集

研究者コミュニティーに開かれた国語研2

- ①新領域指定型(公募型)プロジェクトの成果から
- ②研究情報発信センターによる調査データの公開

研究者紹介 著書紹介



AI は皮肉を理解する?

日本語の間接発話理解:

第一言語、第二言語、人工知能における習得メカニズムの認知科学的比較研究

章

皮肉がわかるAIが求められている

皮肉とは、現実に反することを意 図的に言うことで、話し手の否定的 な態度を間接的に伝える表現である。 例えば、いつも時間にルーズで、今 日も待ち合わせの時間に遅れてきた 友人に対する「今日も時間ぴったり だね」という発話を考えてみるとよ い。友人どうしの会話であれば軽い 冗談として受け流すことも可能であ るが、それでも相手が快く思ってい ないことは伝わるであろう。私たち が日常的に会話を交わす中で、あま り皮肉を言う/言われることはない。 一日に何回も皮肉を言う皮肉屋さん とは、あまり付き合いたくないのが 人情である。では、なぜそんなニッ チな表現をAIに理解させる必要があ るのだろうか。

2014年6月に米国のシークレット サービス(要人の警護や捜査・諜報 を行う機関)がソーシャルメディア における皮肉的な発言を検出するソ フトウェアを求めているという ニュースが報じられた。シークレッ トサービスは、犯罪やテロなどの予 兆を把握するために、常時Twitterな どのソーシャルメディアに投稿され る発言を注視している。その際に、 センチメント分析という手法を用い

て、発言の極性(発言内容がポジティ ブかネガティブか)を自動で判断し ている。しかし、皮肉の場合には(前 述の例のように)表面的にはポジティ ブなことを言いながら、真意はネガ ティブである。よって、通常のセン チメント分析では誤ってポジティブ な内容(=問題ない発言)と判断し てしまうため、その前に皮肉かどう かを検出することが求められるので ある。

AI は教師データから皮肉を理解 する

このようなニーズもあって、自然 言語処理(言語を扱うAIの一分野) では、ソーシャルメディアの発言を 対象とした皮肉検出の研究が盛んに 行われている。そこでは、大量の皮 肉発言と皮肉でない発言を収集して 教師データとし、ディープラーニン グなどの機械学習に基づく手法を適 用して、皮肉かどうかを判定する分 類モデルを自動で学習する。その性 能は、現在では少なくても80%以上 の発言を正しく判定することができ る。

一方で、学習された分類モデルは 一般的に中身はブラックボックスで あって、私たちが理解できる形でど のように皮肉と判断したかを知るの

は容易ではない。また、Twitterでの 皮肉は日常会話の皮肉とは異なる性 質をもっている。例えば、「今日も遅 れて来るなんて、何て時間に正確な んだろう」というように、言語表現 内に矛盾する内容を含めるような表 現が多用される。これは、その場で 発話のやり取りをする日常の会話と は異なり、あくまでも単発の「つぶ やき であるという性質による誤解 を避けるためであろう。(このような 現象は言語学の研究対象として興味 深い。)

人間の皮肉理解をモデル化する

そこで、私たち人間が皮肉をどの ように理解しているかという認知プ ロセスを考えて、それをモデル化す るというアプローチを考えてみる。 皮肉が発話される場面を観察すると、 話し手はあることを期待していなが らそれが満たされておらず、その現 実に対して否定的な態度を持ってい るという特徴がある。このような状 況をアイロニー環境と呼ぼう。アイ ロニー環境を相手に暗黙的に提示す るのが皮肉であり、そのための発話 は期待に言及しつつどこか逸脱を含 むという特徴が見られる。皮肉を言 われた側は、その表現上の手がかり

を参照しつつ、今の状況が皮肉を言いやすい状況かどうかを判断して、 皮肉かどうかを理解している。これが、私が提案している暗黙的提示理 論という枠組みである。

この枠組みをベイズモデルという 統計的な手法を援用してモデル化してみよう(図1)。求めたいのは、ある状況Cで発話Uが皮肉と判断される確率である。この確率は、事前である。こので皮肉が発話される確率(専なで皮肉が発話される確率で皮肉が発話した用いるでときに発話Uを用いるの積でしたの発話Uを用いる。 暗黙的提示理論となるでは状況を開いたがアイロニー環境となる確率を表してがアイロニー環境となるである。大 大度は状況Cでの発話Uが暗黙的提示になる確率を表している。

このモデルを用いて、以下のよう な文章における発話の皮肉確率を推 定する。

状況:あなたは、友だちの引っ越 しの手伝いに行きました。 あなたはすぐに疲れてしま い、みんなの邪魔になって いました。友だちはあなた に言いました。

発話:「手伝ってもらって本当に 助かったわ」

また、同じ文章を読んだときに皮肉であるかどうかを5段階(0~4)で複数の人に尋ねた。40件の文章に対する結果が図2である。全体的に皮肉確率は低く見積もられているが、人間の判断と弱い相関(r=0.39)が見られた。また確率0.1以上で皮肉と判断すると、27件(67.5%)が正解となった。性能としてはまだ不十分であるが、大量の訓練データから分類モデルを学習しなくても、理論に基づくモデル化でこのくらいの判断ができるのである。

AIは皮肉を理解できるか

 (電気通信大学・教授/内海彰)

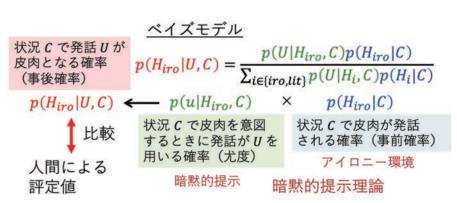


図1 ベイズモデルを援用した暗黙的提示理論の枠組み

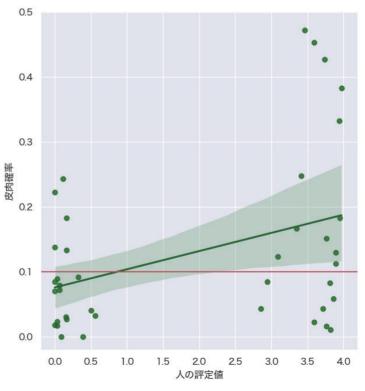


図2 同じ文章を読んだときの皮肉確率と人の評定値

地域社会の言葉は何によって変化するか?

音声の個人内変化を推定するためのコーホート分析の有効性に関する実証研究





1971年



1991年



言葉は変化する

言葉はいつの時代においても変化 します。日本全体としても、特定の地 域社会としても変化します。

たとえば単語レベルで考えると、社会に新たな事物や概念が生じれば、それを表わす言葉も当然必要となり、追加という方向で言葉が変化します。昨年来の新型コロナウイルス」という言葉をはじめとし、関連する語がたくさんとまれました。もし新型コロナウイルス禍がなかったら日本語に追加されなかった語です。逆に、社会から事物を概念が消滅すれば、それを表わす言葉も不必要となり、消滅という方向で言葉が変化します。

このほか、よその言葉の影響を受けて変化が生じることもあります。方言の共通語化などはその典型的な例です。

言葉が自ら変化するわけではない

「言葉は変化する」と言いましたが、 厳密に言えば言葉が自ら変化するわけ ではなく、これまでなかった言葉を使 う人が現われたり、これまで使われて いた言葉を使う人がいなくなったりす るということです。言葉はそれを使う 人間に支えられて存在するものである ことを考えると、言葉の変化というの は、じつは言語使用者の変化というこ とと密接に関係しています。

この観点から、地域社会の言葉は何によって変化しているのかをさらに詳しく見ていきましょう。

地域社会の言葉を変化させるもの

地域社会は当然ながらその構成員により成り立っています。しかしその構成員はずっと固定しているわけではなく、時間の経過とともに入れ替えが生じ、100年も経てばほぼ全員が入れ替わります。転出・転入という形での構成員の入れ替えもあります。

ある言葉の使用にもし一貫した年齢差があるならば、それを使う高年層が構成員からいなくなり、逆にそれを使わない構成員が新たに加わると、その言葉がその地域社会で衰退する方向に変化することになります。

しかしこの考えは、個々の構成員の 言葉は生涯変わらないということを暗 黙のうちに前提としています。個人内 での変化は本当にないのでしょうか。 もしそれがあるとすれば、地域社会の 変化は、構成員の入れ替えに加え、個 人内変化も要因の一つとなっていることになります。ある言葉が社会全体として衰退に向かいつつあるならば、特定の個人もそれを使わなくなるということや、逆にある言葉が普及しつつあれば特定の個人も使い始めるということはおおいに考えられそうです。

コーホート分析の有効性を検証 する

ただ、特定の多数の個人をたとえば20年後に追跡調査するというのは研究としての難度が高く、とりわけ個人レベルでは現実的ではありません。

そこでその代替として、2度の調査におけるコーホート(同時期出生集団)を比較することで、個人レベルで変化が生じているか否かを推定するという方法が用いられます。

しかし、異なる個人の集団を調査するわけですから、本当に個人が変化しているか否かをとらえているか疑問が 残ります。

そこで、国立国語研究所が多数の個人を追跡調査したデータ(パネルデータ)を用い、これとコーホートとを比較することで、コーホート分析の有効性がどれほどであるかを検証しました。もし両者がほぼ一致すれば、難度の高い追跡調査に代えてコーホート

調査をすることで、地域社会の変化 に寄与する個人内変化の度合いを知 ることができそうです。

「鶴岡調査」のガ行鼻音を分析 すると…

国立国語研究所では、方言の共通 語化を把握することを主たる課題とす る調査研究を、山形県鶴岡市におい て、多人数の市民を対象に経年調査 しています。

第3回調査 (1991年実施) では、音声項目にガ行鼻音が追加されました。これは方言の共通語化というよりも、ガ行鼻音を持っている地域ではその衰退が顕著であることから、鶴岡調査でも追加されたものです。具体的には「鏡」のガを調べています。

この項目は20年後の第4回調査 (2011年実施)でも引き続き調査され ています。このデータを用いて、両調 **査のコーホートの変化と、それと出生** 時期が重なるパネル(両調査に回答 者となった人)とを比較しました。た とえば、1950年代生まれで第3回調 杳に回答した集団と第4回調査に回答 した集団を比較するとともに(個人は 異なる)、1950年代生まれで第3回調 査に回答しかつ第4回調査にも回答し た集団を比較し (個人は同一)、この 2つの調査結果を比較するということ です。前者の変化の方向や幅が後者 とほぼ同じであれば、難度の高いパネ ル調査に代えてコーホート調査を行え ばよいということになります。

なお、音声の聞き取りは微妙なところがあり、人によって判定が異なることもありますので、確実なデータを得るためには聞き取り者を同一にする方が安心です。

さいわい、国立国語研究所のオフィシャルデータとは別に、第3回調査と 第4回調査のデータを筆者が一人で聞き取ったデータがありますので、これを用いることにします。

継続調査とパネル調査の比較

ガ行鼻音で発音した人の割合について、第3回調査(1991年)と第4回調査(2011年)の差(「第4回調査の数値一第3回調査の数値」)を、継続調査とパネル調査に分けて示したのが次のグラフです。「0」に近いほど2回の調査の差が小さいこと、つまり変化が小さいことを示しています。プラスはこの20年間にガ行鼻音で発音する人が増えたこと、マイナスはガ行鼻音で発音する人が減ったことを示しています。横軸は回答者の生年です。1930年代生まれから1970年代生まれまで10年刻みとしました。

1970年代生まれのパネル調査の回答者は14人にとどまり、安定した数値でない可能性がありますので、これ以外の生年層に注目して見てみましょう。

青色の点線の継続調査の数値を基準に考えると、赤色の実線のパネル調査の数値は、どの生年層においても1割ほど高い(ガ行鼻音を持っている人の割合が高い)ことが分かります。言葉の調査に複数回協力してくれたパネル調査の回答者は、言語使用や言語意識について平均以上に関心が

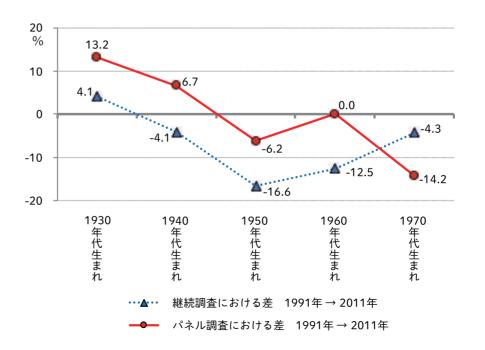
高く、規範意識も強いことが考えられ そうです。そのため、そうではない人 を含む継続調査よりも、ガ行鼻音を保 持する人の割合が高く、このような違 いが生じた可能性が考えられそうです。

つまりパネル調査は、個人の変化の有無を確実に把握できるという点では大変有用ですが、集団の代表性は必ずしも保証されているわけではなく、その点で限界がありそうです。変化の有無や度合いを数値的にとらえるのであれば、集団の代表性が高い継続調査のコーホート分析の方がむしろ有効かもしれません。ただしこのことは現時点では仮説にとどまり、同様のことがガ行鼻音以外でも広く該当するか否かを今後確認する必要があります。

参考文献

尾崎喜光 (2019) 「言語変化を把握するため の継続調査と同一個人追跡調査の関係に 関する研究―国立国語研究所実施の鶴岡 市調査に見るガ行鼻音衰退の事例―」 『計量国語学』 32-2

(ノートルダム清心女子大学・教授/ 尾崎喜光)



特集 研究者コミュニティに開かれた国語研2 ②共同利用型(公募型)プロジェクトの成果から

受付

調査員によって調査の結果はどれほど変わるのか?

第1~3次岡崎敬語調査調査票・調査員記録簿の分析



岡崎敬語調査とは

国語研はこれまでに日本全国各地 で大規模な言語調査を行っています が、岡崎敬語調査はその代表的なも ののひとつです。愛知県岡崎市の住 民を対象にした敬語の使用と意識に 関する調査で、1953年、1972年、 2008年と3回の調査が行われました。 それぞれの回で独自に取ったランダ ムサンプルと、過去の調査の回答者 を再び調査するパネルサンプルの2 種類のサンプルがあります。これに より、岡崎市全体での敬語の変化と ともに、個人がどのように敬語使用 や敬語意識を変化させていくかとい う観点からの分析も可能になってい ます。

岡崎調査でもう一点特記すべきことは、回答がすべて文字起こしされていて、報告書と共にデータベースとして公開されている点です。55年間3回の調査での回答者の発話が、すべて公開されているというのは、非常に画期的なことです。

岡崎敬語調査でわかったこと

その岡崎敬語調査からわかったことは、大まかに言えば、(1)全体とし

て岡崎市の敬語使用がより丁寧な方向に向かっていること、(2)敬語は中年層でもっとも丁寧になり、その後の加齢につれて丁寧さが下降すること、(3)敬語使い分けの基準が社会的立場の差異から心理的立場の差異へと移行し、敬語の民主化・平等化が進行していること、の3点でしょう。

(1)は全調査について、丁寧さの段階付けで得られた敬意の平均値の比較をしてわかったことです。岡崎市民は、明らかに以前より丁寧な話し方をするようになったのです。

(2)は、年齢別の分析から判明しました。いずれの調査でも、社会活躍層と言われる中年あたりで丁寧さは最大値に達し、その後加齢につれて下降するパターンを見せています。高年齢層での下降には、社会との接点が少なくなることが原因として考えられるでしょう。

(3)は話者の属性による差が縮まっていること、それにもかかわらず場面による差が見られるという事実から導かれることです。社会的身分の上下によって使い分けられていた敬語が、自分と話し相手との距離によって使い分けられるように変化しているのです。

岡崎敬語調査への疑問

さて、こうした岡崎敬語調査については、ひとつ疑問があります。この調査には3回の調査を通じています。このべ306人の調査員が参加して、ま方には大学院生もいれば、長年いまでは大学でしてきた研究をしています。では大学の研究をしておかのないが予しています。ではないではないでしょうか。

実はこうした調査員のもたらす効果というのは、社会言語学のみならず、社会学でも盛んに研究されているトピックです。岡崎敬語調査でもこの調査員効果を分析してみようと考え、このプロジェクトを立ち上げた次第です。

調査者個人差の分析とこれから

次のグラフは、第3次調査に参加 した28人の調査員について、質問ご とに得られた丁寧さの段階付けの平 均値をプロットしたものです。縦軸

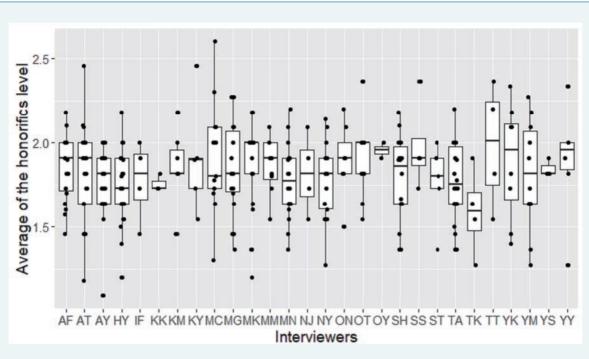


図 調査者ごとの各質問の丁寧さ段階付け平均値の分布(第3次岡崎敬語調査)

は段階付けの平均値を、横軸は調査 者を表します。調査者ごとに11個の 質問のそれぞれの平均を出してプ ロットしているので、各調査者のと ころに11個(=質問の数)の●があ ります (重なっているので11個ない ように見えるところがあります)。同 じ調査者が担当したとしても、質問 による段階付けの平均点の開きには、 かなり大きな差があるようです。上 下にかなり大きく差が出ている調査 者もいれば、小さくまとまっている 人もいます。質問によって段階付け の平均点が異なることは当然として も、調査者による差は無視できるも のなのでしょうか。

この疑問を解くために、データを 多変量解析のひとつである一般化線 形混合モデルという方法を使って分析してみました。その結果、回答者 の性別、年齢、質問、回答者のとを 差を組み入れたモデルがもっると、 でするの段階付けを予測のために、 調査者の個人差は考える必要はの いうことです。一見調査者の個人差 と見えたのは、実は回答者ごとの違いや質問差などによるものだったということになります。

これだけの大規模調査で調査員の 個人差がなかったことを、どう解釈 すればいいでしょうか。まず、各質 問で設定された状況がわかりやすく、 回答者にとって回答しやすいもので あったこと、そして調査マニュアル や調査前の調査員への説明が充実し ていたため、調査員の個人差が生ま れにくかったことが考えられます。 加えて3次調査であることから、調 査の積み重ねによって調査方法が洗 練されてきたこともあるでしょう。 調査において調査員はいわば「黒子」 であり、黒子によって調査結果が変 わることは避けねばなりません。岡 崎敬語調査は、少なくとも3次調査 はその点で成功していたということ ができるのです。

この分析は第3次調査にのみ限定したものでした。では果たして同じ 岡崎敬語調査の第1次、第2次調査は どうだったのか。また鶴岡の共通語 化調査を始めとする他の大規模調査 ではどうだったのか、という新たな

疑問が湧いてきます。こうした疑問は実際にデータを分析しないことには解けません。研究の次の段階として、こうした分析を進めていくことを検討しているところです。

参考文献

井上史雄(編)(2017)『敬語は変わる― 大規模調査から分かる百年の動き―』東京:大修館書店

Matsuda, Kenjiro (2020) "Random Effects in the Third Survey of the Okazaki Survey on Honorifics". Céleste Guillemot et al. (eds.) *ICU Working Papers in Linguistics* 10: 57-64.

(神戸松蔭女子学院大学・教授/ 松田謙次郎)

PROJECT

専門用語を取り込む工夫

難解用語の言語問題に対応する言い換え提案の検証とその応用

田中牧郎

「クラスター」と「オーバーシュート」

コロナ禍の1年半、専門用語起源 の多くの新語が登場し、あるものは 定着し、あるものは淘汰された。そ の定着や淘汰の過程では、新しい概 念の普及に役立った言い換えや言い 添えの工夫もあれば、混乱につなが る不作為も見られた。

「クラスター」は、当初は、わかり にくいと言われたが、「感染者集団 | 「集団感染」などと言い添えが行われ るうちに、すっかり定着してきた。 「オーバーシュート」も、はじめ、「感 染者の爆発的な増加|「爆発的な感染 拡大 などと言い添えを伴いながら もよく使われたが、やがて使われる ことが少なくなっていった。この違 いは何によるのか。

「クラスター」は、上述した言い添 えを伴いつつ、「密」「3密」「濃厚接 触しなど関連する多くの新語ととも に、感染予防に関する具体的でわか りやすい語彙体系を確立し、個々の 語の意味の輪郭も明確になっていっ たと見ることができる。

一方、「オーバーシュート」は、「実 効再生産数 | 「指数関数的 | などの関 連語とともに用いられたものの、そ れらは数学的な思考が求められる抽 象的で難解な単語群であり、一般の 人が理解しやすい語彙体系は形成さ れなかった。

もしも、感染者数の増加の科学的 な見通しを国民で共有することで、 人々に行動変容を促したり、医療体 制の整備を進めたりする政策を取る うとするのであれば、専門家、政府、 報道機関は、こうした用語への言い 添えや説明の仕方に工夫をこらすこ とで理解を得、感染拡大防止への協 力を呼びかけるメッセージを発する べきではなかっただろうか。

「エアロゾル感染」と「マイクロ飛 沫感染」

科学的な内容を含む専門用語の言 い換えや説明は、とても難しい。例 えば、「エアロゾル感染」「マイクロ 飛沫感染」という専門用語は、当初 から注目されていた「飛沫感染」に 加え、後になって報道される機会が 増えてきたものである。飛沫よりも 小さな粒子によって空気中をウイル スが漂うことによる感染を意味する 語だが、飛沫を直接吸い込むことで 感染する「飛沫感染」と、換気され た車内や店内などでも感染する「空 気感染」との間にあって、その位置 づけがわかりにくい。ほぼ同義の二

つの用語を統一する報道がなされた ことがあるが(図1)、現在も二つが 並存したままで、混乱をきたしてい る。語彙体系が形成されないまま放 置されてしまっている。

このように、専門用語を一般の 人々にも理解できるわかりやすい語 **彙体系に位置づけていくことは容易** でない。筆者が国立国語研究所在職 中に担当した、「『外来語』言い換え 提案 | (2002~2006年) と「『病院 の言葉』を分かりやすくする提案 | (2007~2009年) は、そこに挑戦す るプロジェクトだったが、苦心した もののうまくいかなかった言い換え も多かった。当時はうまくいかな かった提案の例から、その後、語彙 体系の形成に向かって状況が改善し てきている事例を紹介しよう。

「モチベーション」と「インセンティブ」

図2のように、「『外来語』言い換 え提案」では、「モチベーション」を 「動機付け」と言い換え、「インセン ティブ」を「意欲刺激」と言い換え ることを提案したが(図中の★の数 は、定着度の段階を示しており、国 民全体では「モチベーション」は下 から2番め、「インセンティブ」は一

印刷版のみ掲載 (PDF版には著作権の都合上、非掲載)

図1 「マイクロ飛沫感染」に用語を統一することを報じる新聞記事 (朝日新聞 2020年7月31日 夕刊 11ページ)

番下の段階)、いずれの言い換え語も あまり使われずに終わった。

「モチベーション」は、2000年代後半から徐々に普及が進み、一般紙での使用頻度は増加していき、現在では、言い添えもなしによく使われており、定着かその寸前まで進んでいると見てよかろう。一方、「インセンティブ」は、使用頻度は増加傾向にあるものの、現在も普及は不十分で、一般紙で使われる場合は、言い添えに「意欲刺激」が使われることはほとんどない。

「『外来語』言い換え提案」では、文脈によっては用いることができる「その他の言い換え語例」として、「モチベーション」には「意欲」「やる気」などを、「インセンティブ」には「誘因」「動機付け」「報奨金」「優遇措置」などを示した。

現在ほぼ定着した「モチベーション」の使用例を観察すると、「モチベーションを上げる」「モチベーショ

ンを保つ」のように、「意欲」「やる 気」の意味で使われている場合がき わめて多く、当初の意味の中心から は少しずれたところに、この語の居 場所は定まったようだ。

「インセンティブ」が一般紙で使われるときに言い添えられる語を調べると、2000年ごろは、経済の専門用語としての意味を示す「誘因」や「報奨金」「優遇措置」などが多かったが、2010年ごろまでには「動機付け」も

多くなり、2020年には「動機付け」が大部分を占めるようになっている。何らかの行動を起こすのに「補助金がインセンティブになる」というように使われるのが典型例で、行動に対する動機付けという意味だ。この意味に落ち着きつつある「インセンティブ」も、一般語としての意味の輪郭が固まりつつあると見ることができるので、今後定着に向かう可能性があるだろう。

人がある行動に出る際、外面から動機付ける「インセンティブ」が働いたり、内面にある意欲の「モチベーション」が活性化したりする、それぞれの局面を指す語が、一般語として体系化されつつあるわけだ。新しい語が定着するとは、語彙体系が形成されることにほかならない。

専門用語の中には、社会の変化に伴って、一般語彙に取り込むことが望まれるものがある。その必要性は、緊急時に痛感されるが、その実現には困難を伴うことも多い。色々な表現で言い換えや言い添えを工夫することを通して、わかりやすい語彙体系が形成されていく。そうした表現の工夫を、平時から様々な立場の人が重ねることで、よりよい日本語の語彙体系をつくりやすい土壌を養っておくことが大切だ。

(明治大学・教授/田中牧郎)

モチベーション

全 体 60 歳以上 ★★☆☆ ★☆☆☆

| 言い換え語 | 動機付け

|その他の言い換え語例 |

意欲 やる気 士気

インセンティブ

全 体 60 歳以上 ★☆☆☆ ★☆☆☆

| 言い換え語 | 意欲刺激

その他の言い換え語例

誘因 動機付け 奨励金 報奨金 優遇措置 意欲

図2 「外来語」言い換え提案 (国立国語研究所「外来語」委員会編『外来語言い換え手引き』 (ぎょうせい、2006年、p.49, p.174)より)

言語学を学ぶことで

得られること

川添 愛 KAWAZOE Ai /作家

「言語学を学んだからといって、言葉がうまく使える ようになるわけではない。物理学を学んだからといっ て野球がうまくなるわけではないのと同じだ」。これは、 私が言語学を専攻していた学生時代、先生方によく聞 かされていた言葉だ。

これが言わんとしているのは、研究対象となる物事についてよく知っていることと、それを日常の中でうまく扱えることとは、基本的に別の話だということである。実際、言語学は言葉の上手な使い方を研究する学問ではない。とくに私が専攻していたのは「現代日本語の文法や意味についての理論的研究」であり、普段の生活の中で言葉を操るコツとか、コミュニケーションのノウハウなんかとは直接関係はない。

とはいえ、周囲を見回してみると、知り合いの言語 学者の中には言葉を巧みに操られる方が非常に多い。 私自身も、そもそも言語学に出会わなかったら、文章 を書く仕事などできなかっただろうと思う。いったい、 言語学のどういう側面が、自分の「言葉の使い方」に プラスに作用したのだろうか。

構造を見極める力

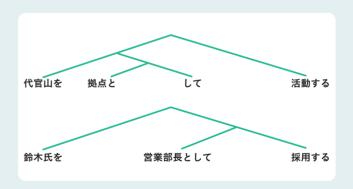
一つには、「文や句の構造を、ある程度意識的に見極められるようになったこと」があると思う。次の二つの表現を見ていただきたい。

- A. 代官山を拠点として活動する
- B. 鈴木氏を営業部長として採用する

これらはどちらも、「~を~として~する」という形をしている。しかし、実は構造が違う。言語学を学んだことがない人に「AとBはどう違うか分かる?」と聞いても、ほとんどの場合「分からない」とか「同じじゃないの?」という答えが返ってくる。そういう人には、次のように「~を~として」の語順をひっくり返してみせると「ああ、なるほど」と納得する。

- A'. 拠点として代官山を活動する
- B'. 営業部長として鈴木氏を採用する

ご覧のように、Aの「~を~として」の部分を「~として~を」に変えたA'は、日本語として不自然になってしまう。これに対し、B'にはとくに問題がない。なぜこういう違いが出るのかというと、次の図のように、AとBでは構造が異なるからである。

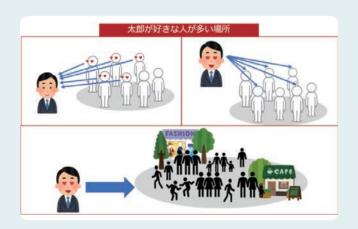


言語学者でなくとも、日本語を理解する人々はみな暗にこういった「構造の違い」を認識しているが、語順を変えるなどといった操作をしない限り、その違いは意識に上ってこない。私自身は、言語学を学んだことで、多少なりともこういった構造を意識に上らせることができるようになったと思う。

曖昧さを分析する力

また別の側面として、「曖昧さとその要因に敏感になったこと」が挙げられると思う。たとえば、「太郎が好きな人が多い場所」という表現は曖昧だが、実際に何通りぐらいの解釈が可能か考えてみていただきたい。

まず、「太郎が好きな」という部分が「人」のみを修飾しているのか、あるいは「人が多い場所」全体を修飾しているのかという曖昧さがある。前者の場合は「太郎が好きな人が、たくさん集まっている場所」という意味になり、後者の場合は「太郎が好む、人の多い場所」という意味になる。さらに前者は「太郎が好きな人」の部分が「太郎のことが好きな人」なのか、「太郎が好いている人」なのかで曖昧になる。つまり、少なくとも次の図のような三つの解釈があることになる。



ただし、「太郎が好きな人が多い場所」の表現の解釈はこれらにとどまらない。たとえば「太郎が好む、人の多い場所」という解釈には、さらに「人が多い場所のうち、太郎が好む場所」なのか「人が多い場所という、太郎が好むもの」なのかといった曖昧さもある。また、上記のどの解釈にも、句全体が「どこか具体的な場所を指しているのか、そうでないのか」という、また別の次元の曖昧さがある。たとえば、「私は今日、太郎が好きな人が多い場所に行ってきた」と言えば、「太

郎が好きな人が多い場所」はどこか具体的な場所を指している。これに対し、「どこでもいいから、太郎が好きな人が多い場所に行ってみたい」などと言えば、件の表現は具体的な場所を指しているのではなく、場所の特徴とか属性のようなものを表していることになる。

別にこの例が特別なわけではなく、私たちが普段使う言葉のほとんどは膨大な曖昧さを孕んでいる。たいていの場合、私たちは文脈や常識を使って曖昧さを排除できるが、たまにそれがうまくいかず、誤解をしたりされたりすることがある。言語学を学んだからといってそういった「事故」を完全に防げるわけではないが、たとえ後からでも曖昧さの原因を分析できるというのは、言葉を適切に使う上で重要だと思うのだ。

言語学は、外部の人から見れば、何をやっているのか理解しにくい分野だ。実際、言語学の外に出てみれば、表面的なところだけを見て「言語学者は誰も言わないような変な文ばかり観察している」とか、「言語学はリアルな実例を見ないから役に立たない」などと誤解している人もいる。あまり学問について「役に立つ」とか「立たない」とか言いたくないが、それでも「理解しにくい」ことを「役に立たない」と同一視されるのは不本意だ。言語学の知見の大切さを広く伝えるにはどうしたらいいか、日々考えを巡らせている。



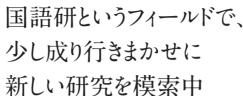
かわぞえ あい ●九州大学文学部、同大学院ほかで理論言語学を専攻し博士号を取得。津田塾大学特任准教授、国立情報学研究所特任准教授などを経て、言語学や情報科学などをテーマに著作活動を行う。著書に『ふだん使いの言語学』(新潮社)、『ヒトの言葉 機械の言葉』(角川新書)など



研究者紹介 021

窪田 悠介

理論·対照研究領域 准教授



くぼた ゆうすけ●1980年神奈川県出身。オハ イオ州立大学でPh.D取得 (2010年)。日本学 術振興会PD、同海外特別研究員、筑波大学助教 を経て2019年4月より現職。専門は統語論・意 味論・計算言語学。好きな言語はRuby。

-- 研究者になったきっかけ

自分でもよくわかりません。ただ、何 か心にひっかかるものがあったせいでこ こまで続けてこられたのだと思うので、 きっかけは生成文法に初めて接したとき に覚えた違和感のようなものかもしれま せん。高校生の頃は英語の文法と数学 が好きでした。大学に入って言語学の授 業を取ったとき、句構造規則を書いて文 の構造を樹形図で分析する方法がある というのを知り、統語論に一番強く興味 を持ちました。高校で習う英文法や日 本語の古文の文法などでは、文法の体系 は分析対象である自然言語自体を使っ て説明されています。そのため、数学の 定義のような感じで、解釈者の直観の余 地が入らない形で文法の分析ができる点 に非常に強く興味をひかれました。

その後、専門に進んで牛成文法の統語 論の基礎を一通り勉強しました。ところ が、ここでの生成文法の第一印象が非常 に悪かった。ちょっと言い方が悪いです が、普遍文法とか言語の生得性とか大風 呂敷を広げているわりには、どうにもい んちきくさい理論で嫌だなあという印象 を持ちました。しかも数学を使って厳密 にやるはずだったのに、専門的な話にな ればなるほど厳密さが失われてふわふわ とした直感的な議論になる。そこに強烈 に違和感を持ちました。そこらへんがお そらく原点です。記述文法や形式意味 論や認知言語学は面白いのに、自分が一 番好きだったはずの統語論が、本格的に やってみようとすると嘘くさいものに見 えてしまってさっぱり面白くない。おか しいなと思って、その後、もっといいも のがあるはずだといろいろ探し回ってい た、というのが現在に至るまでの20年

です。結局最後は自分で自分の理論を 作る羽目になりました。そんなふうに回 り道をしてきて、最近ようやく生成文法 の良さも少しは分かるようになった気が

一 これまでのご研究について

ここ10年くらいは、カテゴリ文法と 呼ばれる、数学と言語学を足して二で 割ったようなものの研究をしてきました。 カテゴリ文法はもともと論理学のほうか ら出てきた手法で、これを言語理論とし て研究する人はほとんどいませんでした。 大学院生時代に最初に見たときは、 ちょっと数学っぽすぎる理論なのでどう かなとも思っていたのですが、試しにと 思って、複合動詞の統語構造や、「太郎 が花子に、みたいな構成素にならない単 語列がまとまりとして振る舞う現象な ど、普通の統語論でうまく分析できない 現象をカテゴリ文法で分析する論文をい くつか書いてみました。

自分でそこそこ納得がいく結果が得ら れたので、理論の基礎的な部分を詰める 作業を行って、最終的に、カテゴリ文法 の研究の流れの中から出てきたアイデア をいくつか組み合わせる形で新しい言語 理論として提案する、というのを博士論 文の研究で行いました。わりと面白い理 論になったので、その後博論の研究に基 づいて指導教員の一人だったオハイオ州 立大学のRobert Levine 先生と10年くら い共同研究を続けています。去年、成果 を本にまとめて出版して、この仕事に一 段落ついたかなと感じているところです。

ー いま、取り組んでいること

いくつか新しいことに取り組みたいと 思っています。一つは日本語の文法の研 究です。記述文法自体は学生のころか

ら関心を持っているのですが、今までは どちらかというと傍観していました。最 近若手の人たちと共同研究をするきっか けがあったので、少し本格的に取り組み 始めているところです。

もう一つは自然言語処理の分野で言 語学に興味を持っている人たちとの共同 研究です。自然言語処理の世界は今、深 層学習の嵐が吹き荒れていますが、新し い技術をうまく言語学の理論研究を科 学研究として進めるために使う方法はな かろうかと考えています。このあたりは、 海外の学会では、手探り感にあふれる研 究が理論言語学プロパーの学会でも出 始めています。これについてもとりあえ ずは傍観しているところなのですが、あ まりもたもたしていると乗り遅れるかな と思い始めているところです。

- これからしてみたいこと

ここ10年くらいは理論研究の基礎的 なことを集中的にやっていました。その 仕事が一段落したので、少し成り行きに 任せて自分がどういう方向に進むのか様 子を見てみたいと思っています。日本語 の文法や自然言語処理の人たちとの共 同研究のほかに、理論言語学の研究成果 を可視化するデータベースを作るプロ ジェクトに加わったりもしています。特 定の理論での個別の現象の研究と、理論 言語学という分野全体のことを考えるの と、どちらにも興味があるので、しばら くは二つの軸を行ったり来たりしながら 次にどこに狙いを定めるか考えようかと 思っています。





小磯 花絵

副所長,音声言語研究領域 教授

様々なコーパスが切り拓く 研究の可能性を追い求めて こいそ はなえ ●1971年東京都出身。千葉大学・同大学院修士課程で認知科学を学んだのち、 ATR知能映像通信研究所で研修研究員をしながら奈良先端科学技術大学院大学で博士号を取得。 1998年10月より現職。

-- 研究者になったきっかけ

人間の言語行動の背後にあるメカニズ ムを知りたいと思って、最初は心理学を 勉強していたんです。でも実験中心のア プローチにあまり馴染めませんでした。 それで哲学の研究室に出入りするように なり、その時に「コーパス」というもの の存在を哲学の先生に教えてもらいまし た。British National Corpus (BNC) という1億語からなるコーパスがイギリ スで作られていて、その中間レポートが あるから読んでごらんって渡されたんで す。言語行動って色々な要因が関わっ てきますよね。そうした要因を大量の データからボトムアップに探ったり、仮 説を立てて検証したり、シミュレーショ ンしたりするのに、コーパスは最適だな と思い、とても可能性を感じました。

いま思えば、これが研究者の道を歩むことになった最初のきっかけだと思います。「いま思えば」と言ったのは、「研究者になるぞ」と強く思ったことはなくて、色々な人や興味あるテーマに出会い、夢中になって研究に取り組んでいるうちに、気付いたらこの道に進んでいた、という感じです。

これまでのご研究について

修士課程で、言語学や心理学、認知科学、情報工学などさまざまな分野の先生や企業の研究者の方と交流する機会がありました。その活動の中で、みんなで使える対話コーパスを作ってみようという流れになり、エディンバラ大学のMap Task Corpusを参考にコーパスを作ることになったんです。当時、日本語のコーパスはほとんどなくて、コーパスの構築と共有化が重要視されていた頃でした。対話をどのように収録するか、転

記・アノテーションの基準やツールをどうするか、これらの情報をコーパスとしてどう表現するかなど、分からないことばかり。多くの方と一緒に手探り状態でコーパスを作りました。

学生時代はコーパスを活用して話者交 代が円滑に行われる仕組みの研究などを していました。形態論・音声・韻律・相 互行為情報などのアノテーションを行 い、話者交替に関わる特徴をボトムアッ プに調べ、それに基づきモデルを作った 上で、シミュレーションを通してモデル の妥当性を検証する、そんな内容です。 コーパスの存在を知った時に、こんなこ とができるんじゃないかなと、漠然と考 えていたことをなんとか形にすることが できました。ただ、かなりまじめに取り 組んでいましたが、分析できたのはたっ た2時間の会話、話者にして16人。とに かくアノテーションに膨大な時間がか かったんです。アノテーションが豊富に ついたコーパスが共有化されていれば、 色々な研究ができたのに。当時、そんな 思いがありました。

その後、国語研究所に入ってすぐに始まったのが『日本語話し言葉コーパス』(CSJ)のプロジェクトでした。ご縁というのは面白いもので、豊富なアノテーション付きのコーパスを構築・共有化するというプロジェクトに携わることができたんです。さらに1億語規模の書き言葉コーパスBCCWJの構築にも関わりました。BNCを知ってから15年以上が経過。ようやくここまで来たな、という思いでした。

― いま、取り組んでいること

独話中心のCSJ、書き言葉対象の BCCWJと、もともと持っていた関心か ら少しずつ離れてしまったんですが、2016年度には、学生時代に必要性を感じていた会話コーパスの構築・共有化に向けたプロジェクトを始めることができました。リアルな日常生活における多様な場面・話者による会話を対象とする、映像付きの大規模コーパス(CEJC)です。CSJ、BCCWJの構築経験やその過程で築いた人脈があったからこそ、こうした挑戦的なコーパスの構築に挑むことができたと思っています。

今は、完成間近のCEJCを使って、時にはCSJの独話やBCCWJの書き言葉と比較しながら、日常会話におけることばの特徴を実証的に探る研究を進めています。今年度末には本公開となります。是非、多くの方に活用していただきたいと思っています。

これからしてみたいこと

学生時代にやっていたような、会話における人間の行動モデルをコーパスベースで研究したいと思っています。学生の頃はたったの2時間しか分析できませんでしたが、CEJCを活用することで、多様な場面、多様な話者を対象に分析できるようになります。

また、これまでさまざまな種類のコーパスの構築に携わってきて、コーパスが切り開く研究の可能性を見てきました。こうした経験を活かし、まだ十分に共有化が進んでいない分野のコーパスの構築にも携わって行きたいと思っています。





山田 真寬

言語変異研究領域 准教授

ことばの復興を、誰でも、どこでも、 みんなでおもしろがってやろう!

やまだ まさひろ ●1982年北海道函館市生まれ。 国際基督教大学卒業後、米国デラウェア大学に 進学し2010年にPh.D. 取得。京都大学、広島大 学、立命館大学での任期付き研究職を経て2018 年から現職。

― 研究者になったきっかけ

心理学を学ぼうと思って入学した大 学で(間違えて)とんでもなくおもしろ い言語学の授業を受講してしまい、それ から卒業まで、徹夜で宿題をやったり生 成文法という分野の論文を読みまくった り、また似たような友だちにも恵まれ、 彼らと一緒に勉強会をしたり東京中の 大学院に聴講しに行っていました。そし てあるとき、アメリカの大学院にいた先 **輩に「アメリカに来たらそれやってれば** 給料もらえるよ」と言われたのが研究を 仕事にしようと思った最初のきっかけで す。授業料も免除な代わりに、能力がな いと判断されれば即クビというのも、辞 め時がわかりやすくていいと思いました。 卒業後はアメリカの大学院に行くかフ リーダイビングの選手になるかの二択で したが、デラウェア大学からオファーを もらったので研究の道に進みました。

--- これまでのご研究について

アメリカでは主に、論理式と関数を使って人間が言語の意味を解釈するしくみを研究していました。それと、デラウェア大学はフィールドワークと理論研究の両方をやる人が多く、僕もフィールドメソッドという授業を毎年履修していました。この授業には未記述の言語の母語話者が来て、1年かけてその人の言語を体系的に記述します。でも、いつかこういう疑似フィールドワークではなく、身一つでジャングルに分け入って未記述の言語をこの手で…ということを夢見ていました。

2010年度の学振PD (日本学術振興会の特別研究員) に採用されたので博士論文を書いて帰国したのですが、期間中の半分は日本にいないといけないという

学振の制約が。アメリカに1年半戻る計 画がすでにあったので、ジャングルには 行けない…。しかし受入教員の先生が 「日本にも君がひとこともわからん言語 があるんやで」と、宮古島に連れて行っ てくださり、さらにPD同期の友だちに 「宮古島より記述が進んでいない与那国 島に行こう」と誘われ、2010年から与 那国語のフィールドワークを始めました (与那国島は亜熱帯気候なので、ジャン グルでした!)。それから数年、彼と一 緒に与那国語の文法記述をしてきまし た。また研究者仲間に誘われて2015年 からは沖永良部島でもフィールドワーク を始めたこともあり、言語の継承保存も 研究テーマにすることになりました。は じめて琉球語の絵本(『みちゃぬ ふい』 言語復興の港 2016年) をつくったのも この頃です。

― いま、取り組んでいること

数か月ごとに数週間のフィールドワー クに通って島の友だちが増えてきた頃、 「島のことば残せるんなら残したい、で もどうすればいいのかわからん」という 声をよく聞くようになりました。研究者 は体系的な記述を残すことはできても、 談話資料などの言語記録は1人だけでは 残せる量が限られていて、さらに言語の 継承保存(話者数の維持)は言語コミュ ニティの主体的・継続的な取り組みがな くてはできません。でもこれぜんぶやっ た方がいいんだよな…と思い、消滅危機 言語の文法記述・記録保存・継承保存を 並行して行う「言語復興の港」プロジェ クトを進めています。言語の研究者が文 法記述を進め、研究者と言語コミュニ ティが一緒になって言語の記録を蓄積 し、これらをベースにした言語継承を言

語コミュニティが主体的に継続できるようにするプロジェクトです。

主にフィールドワークをしている沖永 良部島の人たちと一緒に、言語記録のト レーニング機会、絵本や動画など楽しく 島のことばを身につけられるコンテンツ、 継承保存のモチベーションになるイベン トなんかをつくっています。ほんとうに たくさんの島の人たちのほか、言語研究 者、アーティスト、デザイナー、役場や 教育現場の先生たちなど、「これぞ協働」 という感じでプロジェクトを進められて います。

- これからしてみたいこと

島のことばが「話せる」ことと「教え られる」ことが違うと理解した人たちが 増え、僕たちが「潜在話者」と呼んでい る、流暢には話せないけど聞いて理解で きる人たちの中からは、島のことばを話 すようになった人たちも。彼らは子育て 世代であり、沖永良部島のことばは復興 を始めたと言えます。沖永良部島の人 たちのサポートはもちろん続けますが、 これは「僕たちが」「沖永良部島で」だ けできたことなのか、「誰でも」「どこで も」できることなのか…。後者になるよ うな普遍知を抽出して、一人ひとりが大 切に思う地域言語がなくならないように したいです。「みんなおもしろがってやっ てる状態」をつくればいいだけなんです けどね。

Book Review *****

日本語研究と言語理論から見た言語類型論

窪薗晴夫, 野田尚史, プラシャント・パルデシ, 松本曜 編



語類型論という分野は自分が手掛けるには敷居が高いが、(I)「言語類型論の研究が個別言語の研究にどのように関わるか」あるいは(II)「類型論研究が理論言語学研究とどのような接点を持ちうるか」には大いに関心がある、という研究者は多いと思われる。本書は、そのような研究者にとって非常に有益な論文集である。

本書の「第I部 日本語と言語類型論」 (第1~8章)は(I)の問いに関わり、日本語の「アクセント」「諸方言のイントネーション」「オノマトペ」「主題」「とりたて表現」「語順」「有対動詞」「移動事象表現」の、いずれも、日本語から言語類型論研究に貢献できる可能性の高いト ピックが選ばれている。

「第 II 部 言語理論と言語類型論」(第9~12章)は(II)の問いに関わり、「最適性理論」「生成文法」「計算言語学」「認知言語学」の4つの理論言語学分野が言語類型論と対置されており、読み応えのあるレビューと評言が提示されている。

国語研は外部の共同研究者の協力が不可欠であるが、本書も、オノマトペ(第3章)、語順(第6章)、認知言語学(第12章)など、このテーマならこの人という人選の妙が随所に光る好著である。

▶堀江 薫 (名古屋大学)

顕在化する多言語社会日本

多言語状況の的確な把握と 理解のために

福永由佳 編 庄司博史 監修



本社会がこれまで持ち合わせてきた「多様性」が、多くの人に認識されるようになりました。近年、これは「ダイバーシティ」として、さまざまなセクシャリティ、人種、年齢、国籍、宗教、生涯などの違いを受用する社会の実現を目指した概念として使われるようになりました。このダイバーシティを論じる上で、重要になるものの一つが言語です。

近年、日本に滞在する外国人の増加による「国際化」「多言語化」が取り上げられるようになりました。日本社会は昔から多言語社会でしたから、「多言語社会日本」に新たな展開が見られたわけです。本書では、多言語社会としての日本をより的確に捉える

視点を設定することを目的に、この分野の研究を牽引してきた研究者たちによる考察がなされています。そこでは多言語社会をめぐる理論的枠組みや具体的な事例に基づいた課題が取り上げられています。

ダイバーシティの観点から見た日本語研究は、今後の言語研究において最も重要な研究課題の一つになるでしょう。多言語社会の「顕在化」はこれからも進んでいきます。その状況をより的確に捉えるための言語研究を行うことを本書では主張しています。その現状、研究の観点を把握する意味でも、必読の一冊です。

▶朝日祥之(国立国語研究所)

コーパスと近代日本語書き言葉の一人称代名詞の研究 近藤明日子

勉誠出版



一つの点で画期的な本だと思う。先行 一研究のきわめて多い、近代語の一人 称代名詞について、知られていなかった数々 の言語事実を明らかにし、その体系変化を 明解に説明した点と、日本語の歴史の研究 において、コーパスを作って研究をするこ とが、いかに大きな力を発揮するかを示し た点とである。

第一の点は、著者の様々な工夫の賜物だが、従来用いられることのなかった文語体書き言葉も含めた、多様な文体の資料を時系列に沿って調査したことと、性差の視点を重視したこととが、特にものを言っている。書き言葉と話し言葉が相互に絡み合いながら進む文体変革と、社会の近代化にと

もなって変容するジェンダー意識とが、一 人称代名詞の体系形成に影響を与えたこと が、総合的にとらえられており、日本語史 研究に新しい視界を切りひらいた。

第二の点では、研究目的に応じてコーパスを設計し、得られたデータに対して、細心の注意と粘り強い思考に基づいて、統計処理や文脈分析を施しているところが、特に読み応えがある。手間をかけ、時間をかけ、データを作り込み、データと対話することで、質の高い研究が成るわけである。コーパスの導入は、日本語史研究を格段に発展させている。

▶田中牧郎 (明治大学)

編集後記

『ことばの波止場』の第十号です。

表紙は「ニホンゴ探検2021」の紹 介画像です。ニホンゴ探検は、子ど もたちが国語研で「1日研究員」とな り、クイズやミニ講義を通じて、こと

ばの不思議に触れていくというものです。子どもだけ でなく大人も楽しめる一般公開イベントですが、新型 コロナ対策で昨年と今年はオンラインで開催しました。

今回の特集は「研究者コミュニティに開かれた国語 研2」で、国語研が所外の研究者から研究計画を公募 して実施された研究プロジェクトに関するものです。ま ず「新領域創出型(公募型)プロジェクト」のうち、1 つのプロジェクトについて成果を紹介しました。次に 「共同利用型(公募型)プロジェクト」のうち、3つの プロジェクトを取り上げました。これらの記事は、前 回の第九号と同じく、研究者コミュニティに利用して いただく大学共同利用機関としての国語研の役割の一 端を紹介するものです。公募型の研究プロジェクトは さまざまな分野をカバーしていることを実感していた だけると幸いです。

コラムは、7月に『言語学バーリ・トゥード: Round 1 AIは「絶対に押すなよ」を理解できるか』(東京大 学出版会) の話題作も刊行され、今、もっとも言語学 の魅力を素敵に発信されている川添愛先生にご寄稿い ただきました。

研究者紹介は、窪田先生、小磯先生、山田先生の3 名を取り上げました。また、書評は3点の研究書を紹 介しました。

今後も国語研の活動を広く知っていただくために、本 誌の充実をはかる工夫をかさねていきたいと考えてい ます。

(横山詔一)

※次号 (vol.11) は2022年3月頃発行予定です。

国語研 ことばの波止場 vol.10

2021年9月30日発行

編集 国立国語研究所研究情報誌編集委員会

柏野和佳子(委員長) 井上文子 五十嵐陽介

福永由佳 横山詔一 松本曜

発行 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 〒190-8561

東京都立川市緑町10-2

電話0570-08-8595(ナビダイヤル)

協力 くろしお出版 デザイン 黒岩二三[Fomalhaut]

https://www.ninjal.ac.jp/ ©National Institute for Japanese Language and Linguistics