

第10回 国語分科会言語資源小委員会(Web開催)議事録

令和 7 年 12 月 23 日
15 時 00 分 ~ 17 時 00 分
文部科学省15階 15F特別会議室

〔出席者〕

(言語資源小委員会委員)相澤主査、石川副主査、神永、齊藤(美)、武田、前川、
森各委員(計7名)

(ローマ字小委員会委員)森山主査、滝浦副主査、大島、川口、川辺、木村、齋藤(純)、
棚橋、常盤、中江、長岡、成川、古田、前田、村上、山本(真)、山本(玲)各委員
(計17名)

(ゲスト)小椋秀樹立命館大学教授、金愛蘭日本大学准教授

(文部科学省・文化庁)山田国語課長、中田課長補佐、武田主任国語調査官、
鈴木国語調査官、町田国語調査官ほか関係官

※ 相澤主査、石川副主査、神永、武田各委員、森山主査、成川、
前田各委員及び鈴木国語調査官を除く事務局は、15F特別会議室にて
参加。

〔配布資料〕

- 1 第9回国語分科会言語資源小委員会議事録(案)(委員限り)
- 2 コーパスを活用した外来語表記のゆれに関する調査(小椋秀樹氏御提出資料)
- 3 新聞経年コーパスを用いた外来語の基本語化の研究(金愛蘭氏御提出資料)

〔参考資料〕

- 1 言語資源小委員会における主な審議事項及び当面の予定(案)
- 2 今後における日本語のデジタル言語資源の整備・活用の在り方(報告)
(令和7年3月17日 文化審議会国語分科会)
- 3 ローマ字のつづり方(令和7年内閣告示第4号)

〔経過概要〕

- 1 事務局から定足数の確認があった。
- 2 事務局から配布資料の確認が行われた。
- 3 前回の議事録(案)が確認された。
- 4 山田国語課長から参考資料3「ローマ字のつづり方(令和7年内閣告示第4号)」に基づいて「ローマ字のつづり方」が内閣告示となったことの報告があった。
- 5 小椋秀樹立命館大学教授から配布資料2「コーパスを活用した外来語表記のゆれに関する調査」に基づいて説明があり、質疑応答が行われた。
- 6 金愛蘭日本大学准教授から配布資料3「新聞経年コーパスを用いた外来語の基本語化の研究」に基づいて説明があり、質疑応答が行われた。
- 7 二つの説明を踏まえ、外来語をめぐる意見交換が行われた。
- 8 次回の言語資源小委員会について、令和8年1月27日(火)15時から予定していることが確認された。
- 9 質疑応答及び意見交換における各委員の発言等は次のとおりである。

○相澤主査

定刻となりましたので、ただ今から第10回、今期4回目の文化審議会国語分科会言語資源小委員会を開催いたします。皆様には年末迫り御多忙のところ御出席を賜りまして、ありがとうございます。厚く御礼申し上げます。

今回も文部科学省の会議室にお集まりいただきつつ、オンラインでも参加いただく形となっております。また、傍聴者の方々もオンラインでこの会議を御覧になられているということで御承知おきください。

なお、本日は非常にスペシャルな回となっております、ローマ字小委員会の委員の皆様も出席されております。本当に、お忙しい中、ありがとうございます。

本日の議事内容が、来期以降国語分科会で検討する可能性がある外来語に関する課題を取り上げているということから、御出席いただけることになったと伺っております。

さて、本日は議事次第のとおり、「言語資源の整備と活用」について、「有識者による事例等紹介」、「その他」という形で協議を行ってまいります。

委員の皆様のお手元には既に事前に資料が届いていることと思いますが、本日は、コ

ーパスを活用して、あるいは御自分でコーパスを構築されて、外来語に関する研究をされていらっしゃるお二人の先生をお招きして、お話を伺います。お一人目は、外来語の表記に関して、立命館大学の小椋秀樹先生、お二人目は、外来語の基本語化について、日本大学の金愛蘭先生にお話をお伺いする予定でございます。その後、意見交換をと思いますので、どうぞよろしくお願いいいたします。

議事に入る前に、事務局から「ローマ字のつづり方」の内閣告示について報告がございます。山田国語課長、よろしくお願いいいたします。

○山田国語課長

国語課長の山田でございます。

「ローマ字のつづり方」につきましては、昨年の5月に大臣から文化審議会に諮問がございまして、その後、主にローマ字小委員会の委員の皆様を中心に精力的に御議論いただいております。そして、今年の7月、国語分科会での御了承を経まして、8月の文化審議会におきまして、見直しの方針について文部科学大臣に答申を頂いております。その後、文化庁の方で、政府としての決定に向けまして政府内の調整を行ってまいりましたけれども、先日その調整が調いまして、昨日、12月22日付けで「ローマ字のつづり方」に関する新たな内閣告示が制定されております。告示の内容につきましては、本日の参考資料3として配布させていただいておりますが、基本的には8月の答申の内容をそのまま反映する形となっております。

今回の、この新たな告示の制定に伴いまして、従前の昭和29年の内閣告示は廃止されております。

答申の取りまとめに御尽力いただきました森山分科会長をはじめ委員の皆様、この場を借りて改めて御礼を申し上げます。どうもありがとうございました。

以上です。

○相澤主査

長きにわたる審議であったと伺っております。本当に皆様の御尽力に敬意を表したいと思います。

ローマ字小委員会の森山主査から何かもし御感想等ございましたら、よろしくお願いいいたします。

○森山主査

この度、大きな変革ということで、委員の皆様、そして事務局の皆様の本当にその尽力のおかげで、新たな日本の表記として、皆さんがより使いやすいものがこのようにできたこと、本当に心から喜ばしく思っております。そして皆様に敬意を表します。また、言語資源小委員会の皆様にも、今後またいろいろな形で御指導をお願いいたします。まずは心よりお礼を申します。ありがとうございます。

○相澤主査

お疲れさまでございました。ありがとうございます。

もし何かこの場を借りてほかに御質問やコメント等ございましたら…。

(→ 挙手なし。)

よろしいでしょうか。

では、議事に入ってまいりたいと思います。まずは、有識者としてお招きいたしました2名の先生方からお話を伺ってまいります。御質問等につきましては、御発表が終わったところで時間を取っていきたいと思います。その後、お二人の先生方の御発表が終わった後の意見交換の時間もございますので、御発言等よろしくをお願いいたします。

まず、立命館大学の小椋秀樹先生から御発表を頂きます。配布資料2に先生の資料がございますので、こちらも御覧いただきながら、よろしくをお願いいたします。お話のタイトルは、「コーパスを活用した外来語表記のゆれに関する調査」と表示されておりますとおりでございます。

令和5年3月に国語分科会は「国語分科会で今後取り組むべき課題」を報告しております。この最初には「ローマ字のつづり方の検討」が取り上げられておりまして、山田課長のお話のとおり、「ローマ字のつづり方」の改定が実現したということでございます。この国語分科会の報告では、平成3年に内閣告示として実施された「外来語の表記」についても、既に30年以上が経過していることを指摘しています。その上で、社会における外来語の書き表し方の状況に変化が生じていないかをよく整理して、もし手当てが必要な状況があるようであれば、見直しを行うことも考えられるということをお話しています。

このような今後の課題があることも踏まえまして、この度、お話をお伺いするということでございます。それでは、小椋先生、よろしくをお願いいたします。

○小椋発表者

立命館大学の小椋でございます。どうぞよろしくお願ひいたします。

現在、立命館大学に勤務しておりますけれども、その前は国立国語研究所に勤務しておりました、今日も御出席の前川委員の下で、「日本語話し言葉コーパス」、そして「現代日本語書き言葉均衡コーパス(BCCWJ)」の構築に携わりました。その中でも、形態論情報報の設計と実装に従事いたしまして、それ以降、コーパスを使った語彙、表記の研究を進めてまいりました。

表記に関しましては、表記の揺れに関心を持っておりまして、その中でも特に外来語の表記について幾つか調査を行いました。今回は、その調査についてお話をさせていただこうと思います。

全体の目次はこのようになっておりまして、最初、揺れの調査(1)というところでは、BCCWJを使った現代語の外来語の表記の揺れに関する調査。そして、もう一つ、調査(2)ですけれども、通時コーパスを活用した外来語の表記の経年変化の調査ということになっております。

まず、「はじめに」というところで、外来語の表記の揺れということでございます。こちら、小学館から発行されています『現代国語例解辞典』の第5版でございます。ここではBCCWJを使ったコラムを設けていらっしゃるしまして、表記の揺れであるとかコロケーションについてコラムが設けられています。その中で、例えば「ディスプレイ」と「ディスプレイ」、「プレー」と「プレイ」とか「メロディー」と「メロディ」のように、外来語の表記の揺れも取り上げられていまして、BCCWJの検索結果を基に、その比率を円グラフでお示しになっております。

今ここで取り上げた「ディスプレイ」と「ディスプレイ」とか「メロディー」と「メロディ」なんですけれども、もしかすると、これは表記の違いではなく、「ディスプレイ」と長音で発音するものと「ディスプレイ」と連母音で発音するもの、そういった発音の違いじゃないかと考える方もいるかもしれません。このことに関しましては、宮島・高木(1984)、こちらは国立国語研究所が1956年発行の雑誌90種を対象に行った語彙調査、そのデータを基にした外来語の表記の揺れについての調査の報告ということなんですけれども、その中でこのように述べられています。「『ピアノ～ピヤノ』は、発音がちがうはずだが、その差は微妙である。『ア』または『ヤ』がかいてあるからといって、そのとおり発音しているとみるのは、あぶない。つまり、これは発音の差のはずだが、じつは表記の差にすぎないかもしれないのである。

(『タイプライタ』とかいている人も、『～ター』とのぼして発音している可能性がないことはない。)それで、発音のゆれとおもわれるものも、純粹に表記上のゆれと区別せずに、とりあげることにした」と書かれております。私も研究の際、この立場に立って、ちょっと広めにそういった現象を取り上げています。

それでは、最初の揺れの調査(1)の方について話を進めてまいりたいと思います。これは「現代日本語書き言葉均衡コーパス(BCCWJ)」を資料として調査を行ったものでございます。まず、BCCWJの中のコアデータ110万語分を使った外来語表記の揺れの概要です。そして、その中で、更に対象を絞りまして、語末長音-「コンピューター」か「コンピューター」か-そういった問題、さらに、表記だけではなくて、「コンピューター」、「コンピューター」という、そういった延ばすか、短く発音するかということで、「日本語話し言葉コーパス」も使った、そういった発音の揺れとの比較ということも行っております。「日本語話し言葉コーパス」の方は、国立国語研究所で作成、構築したモノログのコーパスでございます。

揺れの概要ですけれども、これは110万語分のコアデータを使いまして、どんな揺れがあるのか、見通しを得ようということを目的に行った調査です。ただ、やはり110万という語数ですと規模が小さくて、そういう見通しを得るといった目的を持ってやったのですけれども、どうしても外来語の中では高頻度語を中心に上げてしまったことになっております。後で説明することもあるんですけれども、やはり高頻度語というのは表記が安定しているんで、この表を見ていただくと、揺れを調査しようとしているんですけれども、揺れが余りないような、そんなふうにも見える、そういった結果になっております。ただ、こういった調査を行ったところ、いろんな類型に一応分けてはいるんですが、説明は割愛させていただきますけれども、特にやはり長音に関するところの揺れが大きいのではないかなと思いました。「カテゴリー」か「カテゴリ」かという語末長音、それから「マネージャー」か「マネジャー」かという語中の長音、さらに、「メール」か「メイル」かという長音か連母音かと、こういった表記の揺れというものが多いのではないかと思ひまして、まず最初、語末長音のところを調査していったということでございます。

それがその次のスライドということになります。まず、これ、語末長音の表記の揺れということで、雑誌、書籍、新聞、それから、Web-というのは知恵袋とブログをまとめている、そういったものでございます。こちらの表2は度数2以上のものを語として取り上げています。度数1だと当然揺れませんので、度数2以上ということで、例えば雑誌に関して言いますと、語末に長音を持つ外来語が異なり720語ございます。その中で揺れているものが

136語あった。なお、残りの584語に関しては揺れていないわけですが、どのように揺れていないかという、長音符号を常に付ける形で書かれているものが546語、常に省く形で書かれているものが38語ある—こういった形で見えていくということになります。このように見えていきますと、長音符号による表記が優勢であるということは間違いないかなと思いますが、揺れにはレジスター(媒体)による差異がございまして、Web、書籍というのがやはり揺れがこの中では大きい。新聞はさすがにほとんど揺れない。雑誌はその真ん中辺りなんですけれど、やや書籍、Webに寄っているかなと思います。新聞は各社が表記の基準を設けております。なおかつ、それは国の表記の基準に基づいておりますので、そのところはしっかりと統制されているということかと思えます。Webになると、完全に書き手の自由ということになりますし、書籍も自由度が高いのかなと思います。そういうところがこういった揺れの大きさに出ているように思います。

では具体的にどういう語があるのかということなのですが、表3、表4は、揺れが見られない語をそれぞれレジスター別に度数上位5位まで取り上げております。表3が長音符号を常に付ける形で書かれているもので、割と高頻度語が多いと言えそうですし、また、語末のところを見ると、ア段音のものが多いのかなということが分かります。それから、長音符号を省略する表記のみのもので、新聞は3語しかないので、ここでは除外しております。雑誌、書籍、Webですと、こういった語が挙がるんですけども、表3の語に比べると、やはり頻度が低いということが言えるのと、語末が「ティ」とかイ段音ですかね—そういったものが多いということが指摘できるかと思えます。書籍のところでは「リーダー」と「リーダ」、それからWebも「リーダ」とあるんですけど、こちらの「リーダー」は指導者という意味のリーダーでして、こちらの書籍の「リーダ」、長音符号を省略する表記のみのものは、「カードリーダー」みたいな、そういった「リーダ」ですね。短単位で切ると「カード」と「リーダ」に切れますので、その「リーダ」が切られてここで出てきているということになります。

そして、揺れているものですが、長音符号を付ける表記の割合が40%から60%ぐらいのものというものをここに三つのレジスターで挙げております。これも度数上位五つということで、例えば書籍ですと「パーティー」というのが一番度数が高いんですが、長音符号を付けるものが362であって、省略するものが469である、このような形になっているということでございます。

ここで見てみると、割と語末のところでは何か違いがありそうだなということも出てきましたので、原語の語末で分類して表記の揺れを見てみたものが表6ということになります。

例えば「クリアー」とか「コンピューター」、語末が原語で-arとか-er、そういったものを取り上げていまして、例えば-arですと、異なり語数は45あります。これも度数2以上です。その中で揺れているものが7語あった。そして、揺れていないものは38語ありまして、そのうち、どのように揺れていないかという、長音符号を常に付ける形の表記で出てくるものが38語で、省略するものは出てこなかったと、そのように見ていくということになります。やはり原語の語末で大分揺れの状況が違うということも見えてくるんですが、例えば語末が-dyですね。「メロディー」とかそういった語ですけれども、8割方揺れている。ただし、これは語数が少ないので、ちょっとこのところは何とも言えないところもあるかなと思います。-tyですね。「アクセシビリティ」とか「コミュニティ」とか、そういう語ですけれども、ここは比較的語数が多くて、6割ぐらい揺れているということが分かるかと思えます。基本的には揺れない場合も長音符号を付ける形で定着しているものが多いので、比較的長音符号を付ける表記というものが行われているように思うんですけれども、-ty、これに関しては、省略する形の表記で揺れていないというものが34語ございますので、ういう点ではちょっとほかのものと違うところがあるかなと思います。こういった形で、それぞれ外来語の語末においてこういった違いも見えてくるということでございます。

今問題にした「ディ」か「ディー」か、「ティ」か「ティー」かという問題については、次のスライドの石野(1991)でこのように述べられています。『レモンティー』は、だれもがこのように発音していると思われるが、これを『レモンティ』と書く人がいる。同様に『シティー』と発音している人がほとんどだと思われるのに、広告などでは『シティ』を目にすることがはなはだ多い。昔、暗殺された『ケネディー』大統領の名も、大抵の人が語尾を伸ばして発音していたはずだが、マスコミは統一的に『ケネディ』と書いていた。なぜか世間には、『ティ』や『ディ』の音価を『ティー』『ディー』と誤解している人が多いようである」と述べられています。「ティ」と「ディ」に関しては表記と発音の間にずれがある、そういう指摘かと思えます。

そこで、次のスライドになるんですが、表記の問題だけではなく、話し言葉でどう発音されているかということを見てみようということで、ここではBCCWJの書籍データと、そして「日本語話し言葉コーパス(CSJ)」—これを使いまして、外来語の語末長音を持つもの、それがどう発音されているか、どう表記されているかを調べたということです。これまでと違って延べで出しております。そうすると、表記に関しまして、外来語の語末長音を持つものというのは延べで9万7,583語出てくるんですが、そのうち長音符号を付けない表記というのは1万6,597語、17%ぐらいがそういった形で使われています。一方、話し言葉、

発音の方を見ても、延べで2万1,208語あるんですけども、そのうち語末を伸ばさずに短く発音しているというのは1,630語ということで、7.7%。10ポイントぐらいの開きがある。やはり表記と発音とでどうやらずれが出てきているということは分かるかなと思います。ただし、以下、語末のア段音、イ段音等一母音で分けているんですが、全てのところでそういったずれが生じているというよりは、例えば省略が多いのはやはりア段音、イ段音のところ—そもそも語数が大分違うというところもあるわけですけども、そのが多いということが言えます。特にイ段音ではもう4分の1ぐらいが省略しているということが分かります。それに対して、発音の方もア段音、イ段音のところやはり多いということが分かります。ただ、幾ら高い割合になっても、この差というのは大分大きいかなと思いますので、語末がア段音、イ段音のところずれが大きいということが言えるんじゃないかなと思います。

実際にどんな語が発音で語末を短く発音されているのか見てみたものが表9です。こちらは語末がア段音のものだけを取り上げています。こちら、CSJでは異なりで578語出てきまして、度数50以上の語が50語あります。その中で揺れのあるものは41語です。「ドアー」が88%で一番高いんですけど、これ、見出しの表記を「ドアー」というふうにBCCWJでもCSJでもしてしまっているんです。実際、141例は「ドア」ですので、語彙素(見出し)の表記は「ドア」とすべきだったかもしれません。もし「ドア」という語彙素にすると、この場合、「ドアー」は除外されることになりますので、一番短呼率が高いのは、「トリガー」が「トリガ」というふうに発音される、82.8%ということになります。その次が「クラスター」で、ここでもう一気に比率が落ちます。さらに、次は「センサー」ですけど、さらにここもどんと比率が落ちると、こんな形になっていて、短呼されるものというのはそこまで多いわけではないということも分かるかなと思います。

このように、見てきたように、語末長音ですけども、実際には長音でない、長音符号を付けない表記というのがレジスターによっては割と割合が高かったりする、そういうことが起こっていますし、語末の音によっても、また外来語の原語の語末のつづり、その違いによっても、ちょっといろいろな違いが出てきている。中には省略する表記が多いものがあるということと、それから、表記の方は割とそういった省略ということも行われるところもあるんですけども、発音の方は表記ほど短く発音されるわけではなくて、表記と発音の間にずれが出てきているんじゃないかなと思います。

続いて、連母音の表記です。ここでは原語に[ei]という二重母音を含むもので、「メー

ル]か「メール」か、「プレーヤー」か「プレイヤー」かというところで、この経年変化です。国語研究所の方では通時コーパスも構築、公開されていまして、ここでは「日本語歴史コーパス」、そして「昭和・平成書き言葉コーパス」を使いまして、明治から平成までの変化を見てみたというものです。雑誌データを使っておりまして、基本的には8年間隔で雑誌データは収録されていますので、そういった形で経年変化を見ていくということになります。さらに、発音との比較もするというので、「日本語日常会話コーパス」、こちらの方も後で見せております。こちらの方は、2016年から2019年にかけて構築されたもので、日常生活で自然に生じる会話を中心に記録したもので、200時間のコーパスということになっております。

まず、この[ei]の表記ですけれども、「メール」か「メイール」かと、そういうところ、延べで見えたのが表10でございまして。明治、大正は余り高くない。特に昭和前期も余り高くないんですけれども、昭和後期のところで1割台に達して、平成でも15%ぐらいで、やはり昭和前期・後期の間でちょっと増えているということが分かるんですが、それでも平成になっても15%ぐらいですので、長音表記が主流とは言えるかと思いますが、ただ、これは延べで見ているために、高頻度語の表記が影響していると思われます。

以下、表11に挙げたのは各期の度数上位3語ですけれども、これ見ていただくと、ほとんどの語が長音の表記で安定している。長音表記が100%のものもありますので、こういうところが度数が高いので、この部分が影響して先ほどの割合になっているということが言えるかと思えます。この影響を何とか抜く形で集計できないかということで考えたのがこちらの集計方法でして、例えばこういった表記があったとします。「メーン」が5例で、「メイン」が1例で、「ステーション」が10例で、「テーブル」が2例あったとしますと、先ほどのところだと、延べの語数で集計していくんですけれども、この場合、「メーン」と「ステーション」と「テーブル」ですので、長音表記が異なり3語、それから、「メイン」が出てきますので、連母音表記1語ということで、こういった異なりで見ると、そんな集計をしてみました。この「メーン」と「メイン」というのは同じ語の異なる表記ですので、同じ語をこの場合、二重に数えることにはなるんですけれども、何とかこれで度数の高さ、高頻度語の影響は解消できないかと思ひまして、一応こういった集計をしてみました。その結果が表12でして、明治期で9%ぐらいで、大正になると1割台に達する。やはり昭和後期になると3割近くになって、平成は3割超えている、こういった形で、かなり高い比率、割合になっていることが言えるかなと思ひます。延べで見ても一応比率は高まってきているんですけれ

ども、ちょっと異なりで見た場合、この見方がいいかどうかということの問題はあるかもしれませんが、かなり連母音の表記は増えていっていると考えられるんじゃないかなと思います。

その次に、今度は3期にちょっと区分し直しまして見てみようというのが表13でございませう。この区分は、1期を明治・大正・昭和前期、2期を昭和後期、3期を平成としまして、ここ、明治から昭和前期までかなり大きくくくっているんですけども、先ほども見ていただきましたように、ここは割と語数が少ないので、以下、割合を見ようとしているので、大きくくくってみました。何がしたいかと言うと、語によって、長音の表記で固定してしまっているものもあれば、長音表記が優勢のもの、本当に揺れてしまっているもの、連母音表記が優勢のもの、連母音表記で固定しているものというのがあるわけで、そういった類型で分けてみようと考えました。まずは長音表記率が70%以上の語を「長音」、長音表記率が30%以下を「連母音」で、それ以外のを「ゆれ」と分類する。例えばメールというのが仮に20例あったとして、「メール」という長音表記が15例、「メイル」という連母音表記が5例あった場合、これは長音表記率75%ですので、このメールというのは「長音」と分類する、こんな形で見えていこうというものです。

そこで、まず第3期、平成期の外来語－これは連母音を持つものですが、度数2以上のもので異なり語数で見ていくと、長音が7割弱、連母音が2割強ですね。揺れているものは1割弱ということで、揺れているものは割と少なく、長音が一応主流とは言えるかと思えます。この第3期の度数2以上の語、386語あるわけですが、もちろん、第3期に出現する前に、第1期から出現している語もあれば、第2期から出現している語もありますし、それから、第3期になって出てきた語というものもあります。では、初出時期というのをちょっと探ってみる－初出と言っても、今回の調査の範囲内での初出ということになります－そうすると、例えば「ブレーキ」という語は第1期に初出がありますし、「クリエイティブ」というのは第2期に初出、「リメイク」というのは第3期になって出てきたものです。それぞれの初出時期で、386語のうち、度数2以上のものが1期には95語、2期には116語あるということがこれで分かります。

では、それぞれの初出時期における表記は一体どうなっているのか、そして第3期と比較したらどう変化しているのかということを見ようというのが表14ということになります。そうしますと、第3期に出てきた386語のうち、95語は第1期に初出がある、しかも第1期でも度数2以上の語ということになるんですけども、見てみると、9割が長音の表記という

ことになっています。それから、第2期に出てきたものというものもあるんですけども、これ、第2期になると、長音の比率はぐっと下がって6割ぐらいで、連母音の表記が25%ぐらいとなっています。1期のところを見ますと、1期で長音のものというのは基本的に第3期でも長音になっているということが分かるかと思います。それから、第2期で長音だった語はほとんどが第3期でも長音のまま。それから、第2期で連母音だった語は、やはりそのほとんどが第3期でも連母音ということがここで分かるかなと思います。ということで、結局、ほとんど変化していないということが分かるわけです。語によって表記がもう決まっている、入った時期に表記が決まってしまったというんでしょうか。それで、そのままそれが維持されているという形です。しかしながら、先ほどまで見たように、時代が下るに従って連母音表記の割合が増えていっているんですけども、では、その増えているのは何かといたら、もともと長音だったものが連母音の方に変化したというわけではないだろうと。結局、恐らく古い時代に入ってきて、長母音で入ってきた語がある、長母音で表記された語があるとすると、そこに今度、新たに入ってきた語が連母音の表記のものが多くて、それが積み重なっていく。それが繰り返されていって、結果として連母音表記が増えていったと、そういったことなのかなということが見通しとして立ってきたということでございます。

このことに関連して、荻野(2014)－これはWeb検索で「テークアウト」とか「ブレーキ」という20語を検索して、長音表記が多いのか連母音表記が多いのかを調査した論考なんですけれども－この結果としては、長音表記が圧倒的に多い、それから、長音表記か連母音表記かは語ごとに決まっている、古い時代に日本語に入ってきた語は長音表記される傾向にあると結論付けられていまして、特に2点目、3点目は先ほどの調査の結果と合致するところがあるように思いました。

そして、この連母音表記ですけども、では何で増えてきたのかということで、発音が変わったから連母音表記が増えているのかということで、発音に関して言及している論考としては、先行研究として宮島・高木(1984)ですね。これ、宮島・高木(1984)のところだと、「原音が2重母音という意識があるとき」に連母音表記だと言われています。それから、佐竹(1986)ですと、連母音表記というのは、これは「長音でないという意識が強いことの証明」で、そういう意識が強いというんだったら、長音符号と母音表記との対立は、これはいわゆる表記の揺れじゃなくて、発音の揺れの問題だと述べられています。ここでは「意識」と言われていますので、実際の発音ということを言われているわけではないんですけども、しかしながら、佐竹のところでは「発音のゆれの問題」とも言われていますので、実

際、話し言葉コーパスで、連母音で発音しているか、長音で発音するのか—これを調べてみようということで、調査をしました。

これ、既にやられた調査がありまして、岡田(2004)ですが、「日本語話し言葉コーパス」を使って、原音で[eɪ]がある英略語、外来語の発音の調査をされています。結果としては、外来語では、長母音の[e:]で実現される傾向がある—ほとんど長母音での発音ということが言われています。ただし、「メイン」と「レイン」とか、そういう/ei/+/N/という音節構造の場合には、「メイン」とか「レイン」というふうに連母音で実現される傾向があるということを指摘されています。

私の方では「日本語日常会話コーパス」を使って調べたところ、延べで見ても、異なりで見ても、やはり9割以上が長音表記ということになります。ですので、発音では基本的には長音で「メール」と発音しているし、「プレーヤー」と発音している。ただし、表記では「メイル」とか「プレイヤー」みたいな連母音表記が増えているということで、やはりここでもだんだんそういった連母音表記が増えてきているわけですけれども、結果として発音と表記の間にずれが生じてきているということが確認できたということになるかと思います。

さて、こういった外来語の表記の揺れ、発音の揺れも取り上げてきたんですけれども、その背景ということで見えていくと、1991年に内閣告示、内閣訓令によって実施された国の外来語の表記の基準として「外来語の表記」がございます。ここには、留意事項その1のところでご述べられています。「『ハンカチ』と『ハンケチ』、『グローブ』と『グラブ』のように、語形にゆれのあるものについて、その語形をどちらかに決めようとはしていない。語形やその書き表し方については、慣用が定まっているものはそれによる。分野によって異なる慣用が定まっている場合には、それぞれの慣用によって差し支えない」とありまして、これを見る限り、外来語の表記というのは、語形の統一を図ったり、語形選択の基準を示そうとしたりするものではないということが分かります。また、表記の統一を図ろうとするものでもなく、基本的に各分野の慣用を認めていくというもので、緩やかな性格の基準ということが言えます。

今御紹介した長音に関する基準を見ていくと、長音は原則として長音符号を用いて書くとはあるんですけれども、注2のところでは、例えば先ほどの[eɪ]という連母音では、「エイ」と書く慣用がある場合はそれによるとか、英語の語末の-er、-orなどに当たるものは、原則としてア列の長音として長音符号を使うだけけれども、ただし、慣用に依じて長音符号を省くことができるということで、やはりこういった慣用を認めていくところから

も揺れというものが生じる、そういった要因というはあるんだろうと思います。

以下、これは一つの慣用の例ということでJISの基準を挙げていますが、簡単に触れておきますと、古いJISの基準では、この分野の慣用ということなんでしょうけれども、長音符号を除いたところの拍数で付ける、付けないと、こういった基準がありました。ただし、今のJISの基準だと、主として「外来語の表記」によるとあるので、基本、「外来語の表記」の原則によるという方向に変わってきているのかなと思います。

また、NHKとか読売新聞では、連母音の表記の問題ですけど、元々長音表記だったものが、やはり連母音表記が増えているということを踏まえて基準を変えられている。こういったことも起こっているということになります。

最後、国語施策におけるコーパスの活用についてです。ここに引用した「国語施策は、空気のようなものでなければならない」というのは、今ここにいらっしゃる武田さんの前に主任国語調査官をされていた氏原さんから聞いた言葉でして、やはり国語施策というのは国民に抵抗なく受け入れてもらって、無理なく受け入れてもらって、そして定着していくような、そういった空気のようなものでなければいけないんだということをおっしゃっていました。こうするためには、日本語の今の姿を捉えることが必要なんだろうと思います。そのために、実際に国語課も実態調査をされていて、「改定常用漢字表」のときには漢字出現頻度数調査をされていますし、それから出現文字列頻度数調査もされています。この出現文字列頻度数調査というのは、常用漢字の追加候補になった漢字を真ん中に置いて、その前後1文字ずつの3文字の連鎖を取り出して、頻度順にそれを並べていったものということで、それによって、3文字の連鎖を見ると、一応その候補漢字がどんな単語で、どんな音訓で使われているかということが分かるだろうということで、こういった調査をされています。また、国語研究所からもBCCWJ、これは構築途中でしたけれども、1,730万語分のデータを用いまして、候補漢字につきまして、どういう音訓で、どういう語の表記にどのくらい使われているかと、そういった資料を提供させていただきました。さらに、意識調査としては、「国語に関する世論調査」で、「改定常用漢字表」、それから「ローマ字のつづり方」のところでもそれに関する調査をされていて、意識、実態の両面から日本語の今の姿を捉え、そして必要な改定を進められて、受け入れられる施策というものを実施されているんだろうと思います。

その日本語の今の姿を捉えるということなんですけれども、「改定常用漢字表」答申にはこういうことが書かれています。「現代のような変化の激しい時代にあっては、『言葉に

関する施策』についても、定期的な見直しが必要である」と。そして、中略しまして、「定期的・計画的な漢字使用の実態調査をしていくことが重要である」とあります。ここでは定期的・計画的な漢字使用の実態調査とありますけれども、これは「改定常用漢字表」なので、漢字に限定して書かれていると思います。基本的にこれ、言葉の使用実態、日本語の使用実態を漢字に限らず見ていくということが必要なんじゃないかと思います。それに基づいて定期的な見直しということも考えていくということかかと思ひます。BCCWJは主要なデータは2005年までですので、日本語の今の姿を捉えるのは難しい。授業で使っている、学生も「え？」と言うことがございますので、それはしようがないんですけども、拡張が必要ということになります。そのことを踏まえて、BCCWJ2が今、正に拡張されていて、書籍ですと、2006年から2025年のデータが加えられていくということになります。これができるとやはり、日本語の今の姿を捉えることもできるわけですし、さらに、どのような変化を経て今の姿があるかということも把握できるのではないかと思ひます。これは、そういう点で非常に重要なコーパスと考へます。ただ、BCCWJ2で終わってしまうと、20年後にやはり、日本語の今の姿が捉えられないねという話になってしまいますので、こういったコーパスの構築というのは継続的な拡張をしていって、それによって言語の使用実態の把握が必要ということになってくるんじゃないかなと思ひます。コーパスを作るのは本当に大変ですので、今、前川委員が聞いていらっしゃるの、何てことを言うんだと、大変なことは分かっているだろうとおっしゃるかもしれませんが、私もそのことは本当に身にしみて分かっているんですけども、やはりそういったことが、国語施策のためだけでなく、日本語研究、それから日本語に関する諸科学、それから日本語に関する産業ということに貢献していくデータ、そこでやはり有効に活用されていくデータということが必要になってくるんじゃないかなと考へております。

さらに、コーパスを使うという点では、単語情報の活用ということもやはり必要になってくると思ひます。表記に関する基準ということなので、当然、語がどのように書かれているかということですので、語という観点からの実態調査が必要ということ。国語研究所のコーパスは形態論情報—単語の情報を付与していますので、その中でも特に短単位の情報というものを活用していくということが必要になってくるし、そういうことができるコーパスというのをやはり積極的に使っていただければと思ひます。漢字に関して言うと、どのような音訓で、どのような語(又は語の構成要素)の表記に、どのくらい使われているのかとか、外来語に関して言うと、外来音がどう書かれているのかということに加えて、どのよ

うな語に、どのような表記が、どのくらい使われているのかということ、それを単語という観点からしっかりと見ていくということが必要かなと考えております。

長くなりましたけれども、以上で私の報告は終わらせていただきます。どうもありがとうございました。

○相澤主査

どうもありがとうございました。

では、ただ今のお話について御質問がありましたら、よろしくお願いたします。御意見等につきましては、また改めて最後に時間を設けたいと思いますので、質問中心ということで、何かございますでしょうか。

○成川委員

音引き、長音のありなしのところ、やはりコンピューター関係は、かつては2バイトでも少なくということで音引きを外して、その後マイクロソフトが付けてという、その行ったり来たりなので、どうしてもなるんだろうな。その辺、全く触れていらっしやなかったのは何か理由があるのでしょうか。

○小椋発表者

話の中で事実のところだけ触れさせていただきましたので、そこは割愛させていただいたということでございます。古くは多分マニュアルとか書かれている方の基準でも一大分前ですけどもーやはり長音符号を付けるようにしようということをしていたりとかということで、本当いろいろとそういった、「外来語の表記」というものを踏まえて、原則ということでいこうという話もあったりとか、いろんなそういったところで表記のことについて考えていらっしやるということも存じておりますけれども、今回はそこは割愛させていただいたということでございます。

○成川委員

どうしても多くなるだろうなと思われていて、何も触れられなかったもので、どうしてかなと思った次第です。

○小椋発表者

申し訳ありません、そういうことでございます。

○相澤主査

ありがとうございます。

ほかはいかがでしょうか。オンラインの方も、もしございましたら挙手にてお知らせください。

○石川副主査

では、私から。小椋先生、大変勉強になるお話をありがとうございました。3点ほど質問させていただきます。

まず1点目は今後の対応についてです。全体状況としては長音記号が主である一方、時代が進むにつれて、二重母音表記も増えてきているというお話であったかと存じます。今後についてですが、これらの表記に関して何らかのルールや目安を設けていくべきなのか、あるいは個々人の判断に委ねるべきなのか、お考えがあればお伺いしたいと思います。

2点目は各種要因による選択の影響についてです。同じ書籍でも、内容や分野により、また、想定する読者層ごとに、長音表記か二重母音表記かという選択は変わってくるように考えます。また、単独だと「カテゴリー」と言うが、複合語になると「カテゴリ化」と言いたくなるというように、単独語か複合語かによっても選択が変わりそうです。こうした点は表記選択の傾向を調査する際にどのように考えるべきでしょうか。

最後、3点目は、表記と発音のずれについてです。「冷凍」は「れいとう」と書きますが、実際には「れーとー」と読むように思います。この問題と、外来語表記の問題をどう併せて議論していくのが望ましいのか、お考えを伺えたら幸いです。

○小椋発表者

最初の御質問のところですが、ここは本当に難しいとしか言いようがなく、基本的に、このことについて以前発表したときにも、日本語教育の方からも質問されたことがありまして、どう教えるか非常に悩んでいらっしゃるということは分かるんですけれども、現状からすると、やはり長音でまず教えていくということにならざるを得ないと思うんです。しかしながら、実際には増えていっていますので、このところ、実際どうすべきかという

ことはちょっと、申し訳ないんですけど、実態はこうなっているということは分かるんですけども、なかなか、どうすべきかというところは、今の状況でいいとも言えますし、しかしながら、先ほどの最後の3点目とも関わるんですけども、発音とずれているというところの問題では、結局、原語の知識みたいなところが関わってくることになるので、そこを日本語の外来語表記のところで要求するのかなというところですね。ここはほかの普通の仮名遣いとも関わることではあるんですけど、そのところ、私としては気になることはありません。3点目と少しまとめた感じになりますけれども…。

読み手の問題は多分これは、専門的分野ですと、やはり分野の慣用ということもございまして、大分様相は変わってくるだろうと思います。今回はBCCWJで現代の実態を捉えたいということなので、いわゆるBCCWJのサブコーパスの中でも出版サブコーパスを使用しました。このところでは、もしかしたら割とその中には、出版目録と言いますか、それを基にしてランダムサンプリングしていますので、専門的なものもかなり入っているところはあるかと思えます。ただ、もう一つ、図書館書籍がございまして。あれは図書館に入っているデータ、書籍ですね。それがありますので、そういうものを見ると、状況が少し違ってくるという可能性はあるかなと思えます。ただし、図書館書籍の場合は1986年から2005年までの20年ぐらいの時代幅があるので、現代というよりは、より幅が広がるので、何が要因かがなかなか、ただでさえつかみにくいところがある上に、そのところは、実はそれ、時代の幅、何かそういった時代差ということが利いているんじゃないかということになりかねないので、そこでまだ調査は、今回のところはしていないということになります。

以上でよろしいでしょうか。

○石川副主査

はい。ありがとうございました。

○相澤主査

ありがとうございます。

ほかはいかがでしょうか。よろしいでしょうか。

では、私から一つだけ。本当に素人質問になってしまうかもしれないんですが、外来語そのものの定義についてです。拝見したところは、ほぼ英語由来になっているようなんですけれどもいかがでしょうか。

○小椋発表者

そうですね、基本的にはBCCWJの語種情報を使って調査をしているということになります。その中で、[ei]のところに関しては、基本的には割と英語起源のものに大分限定しながら調査したこともございますけれども、ただやはり、実際やっていると、高頻度語のところなんかに関しては英語起源のものがどうしても多くなるので、この表の中ではそういったものが挙がってきているということで、基本はBCCWJの外来語の語種の情報、それを基にして網羅的にまずは取ってきているということにはなります。

○相澤主査

分かりました。ありがとうございます。

ほかはいかがでしょうか。

(→ 挙手なし。)

小椋先生、どうもありがとうございました。

続きまして、日本大学の金愛蘭先生の御発表に移りたいと思います。金先生の御発表資料は配布資料3となっております。タイトルは「新聞経年コーパスを用いた外来語の基本語化の研究」ということでございます。

外来語が増加していて、人々のコミュニケーションを難しくしている場合がある一方で、外来語や片仮名語がより一般的で身近なものになっているという側面もあります。金先生は御自身で新聞に関するコーパスを構築されて、そのコーパスを用いて、外来語が日本語の中で占めてきた位置の変化を調べられてこられたとお伺いしております。

今回はこの御研究についてお話を伺うということで、金先生、御発表準備をよろしくお願いいたします。

○金発表者

画面も共有されていますでしょうか。

○相澤主査

はい、大丈夫です。よろしくお願いいたします。

○金発表者

ありがとうございます。本日は、オンラインにて失礼いたします。日本大学文理学部の金愛蘭と申します。まず、私の研究を御紹介する機会—このような場を与えてくださり、誠にありがとうございます。いろいろ、まだ分析が足りないところも多々あろうかと思えますけれども、どうぞよろしく願いいたします。

まず、研究タイトルなんですけれども、「新聞経年コーパスを用いた外来語の基本語化の研究」というものです。

本日のお話の趣旨なんですけれども、現在私が取り組んでおります研究テーマとして、20世紀の新聞語彙、つまり書き言葉なんですけれども、この新聞という媒体において、抽象的な意味を表す外来語が基本語化しているという現象が現在起きているということを見つけて、それを、自ら作成した大規模な20世紀後半の経年コーパスに基づいた計量的な調査と、それからケーススタディー—特定の外来語に注目した事例研究—によってそれを実証しようとしているというものです。今日はその要点を簡単にまとめまして、言語コーパスの意義や活用法に関する議論の話題提供ができれば幸いに存じます。

まず、流れとしましては、なぜコーパスを自分で作ろうとしたのかという話。それから、作ったそのコーパスというのはどのようなものだったのかという点。3点目は、いろいろ大変な思いをしながら作ったコーパスからどういったことが分かったのかという点についてお話ししたいと思います。最後に、なぜコーパスが必要なのか、更に言えば、BCCWJのような良質の、とてもいいコーパスが作られることによって期待されることや、どんなことが起こり得るだろうかということについて触れられればよいなと思っております。

では、早速ですけれども、1点目、なぜコーパスを自ら作ろうとしたのかという点です。これは一言で言いますと、やりたいことがあったんだけど、それができるようなコーパスがそれまでにはなかったので作るしかなかった。でも、そのやろうとすることはとても意義のあることで、これは自分にしかできないと思ったということに尽きると思います。

よく耳にすることとして、きっとそうだろうなというようなことが、20世紀の日本語語彙において最も大きな変化の一つである、外来語の増加というのがあるんじゃないかということです。戦後の日本語において外来語というのが徐々に増えてきた、あるいは爆増しただろうというようなことは指摘があるけれども、それを実証したものがなかった。その実証できなかった理由にはいろいろあるけれども、その中に、一番大きな要因としては、それが実証できる電子データというんでしょうか、コーパスというものが無いということです。

この量的な外来語の増加というのは、質的な面においても、つまり基本語彙にも少しず

つ進出するという現象も同時に起きているだろうと言われてしています。実際それについて触れている研究というのがありまして、元所員の宮島達夫氏の研究によりますと、例えば雑誌ですね。雑誌という媒体－雑誌90種を対象とした語彙調査が1956年に公開されているんですけども、それと、40年後にはほぼ同じような、同様の調査方法によって明らかになっている雑誌70誌の語彙調査を比べてみますと、上位1,000語、これがいわゆる一番狭い、基本語と言える部分だと思んですけども、上位1,000語における語種構成、つまり、和語－大和言葉と言われるもの、漢語－漢字でできている中国語由来のもの、あるいはそれを使って作られたもの、字音語と言われるものですね、それから、外来語－この外来語というのは、いわゆる用語ですね。あるいは、近代以降で中国から借りられてきたもの、「ラー油」ですとか「ラーメン」みたいなもの、「イクラ」とかといったものですね。あとは欧米系のもの－いわゆる片仮名語というものです。混種語は、この異なる語種をミックスしているものとなります。これを二つ比べてみた結果がこちらの表になりますけれども、これを見ますと、40年という時代を経て、外来語の比率というのが1.8%から7.1%に激増しているということを宮島達夫氏が指摘しています。この40年の間に雑誌の基本的な語群において外来語の増加は顕著に進んだということがデータに基づいて明らかになっているということです。このような調査は実はそれほど、ありそうでなかったんですけども、これをもっと大規模にできるんじゃないかということが気になったということです。

これは、この90種と70誌における外来語のリストなんですけれども、1,000語の中の18語、71語になります。すみません、1点訂正がございまして、レジユメのところ17語と書いてしまっているんですけども、18語の間違いです。訂正して、おわび申し上げます。90種の外来語のリストを見ますと、こういったものになっておりますけれども、例えば抽象的な意味を表す外来語というのはそれほど数がありませんで、例えば「スポーツ」ですとか「デザイン」辺りになろうかと思えます。それに対して、40年後の70種、70誌の外来語を見ますと、いろいろ増えてきているということが確認できます。例えば、型という意味の「タイプ」ですね。それから、「デザイン」も同じくありまして、「レース」ですとか「モデル」－これは人のモデルの方ではなく、見本とかのような、そっちの、抽象的な意味の方です。「ポイント」ですとか「システム」、ちょっと飛んで「スタイル」ですとか「イメージ」ですとか、この辺りがいろいろ入ってきております。ちょっと後ろのところに行きますと、「ソフトな」みたいなものもありますし、「オープンする」のサ変動詞の「オープン」ですとか、「トップ」、それから「チェックする」の辺り、いろいろ増えているということが確認できます。

重要な点としましては、この抽象的な外来語が基本語彙の中に進出していくというのは、例えば、物を借りてー海外から借用するときに、同時にその言葉は、借りるといった、社会や生活の近代化といったところで起きるものとは違う、例えば具体名詞の場合と違うということがまず一つあります。それはやはり物を借りてくるときの語ではありませんので、言語内的な要因でしか明らかにできません。つまり、従来、和語や漢語の類義語があるのにもかかわらず借用するわけですから、その原因は言語内的でしか追求できないと考えております。

そのためには、語彙表、語彙リストだけでは足りませんで、対象とする資料の原文を入力したコーパスを経年的に作らないといけません。そうすることで外来語とその類義語の意味・用法の変化を同時に調査することで、例えば計量的語彙史研究といったものが必要になってくるというわけです。

2番です。どのようなコーパスを結局作ったのかというところに移ります。

これは、先ほど申し上げましたように、自分でやりたいことがあったんだけど、誰も作ってくれていないし、これは作らないといけないということに陥ったので、そこがスタート地点になりますけれども、修士論文で、外来語の基本語化というところー自分の中では非常に意義のあるもので、明らかにしたいなと思ったんです。「トラブル」という語を取り上げて、ケーススタディーとして始めたんですけれども、これだと、元留学生の私は、博士論文を書いて、国に帰れないんじゃないかという危機感を抱きまして、博士論文では自分で作るつもりでしたということです。ただ、作ろうとしても、技術ですとか知識がありませんので、勉強しながらということだったんです。ありがたいことに、ドクターの3年目に外部資金も、助成も受けながら、また、国立国語研究所というところで任期付きの연구원としてBCCWJを作る作業に少し参加させていただいたということがありまして、そのおかげで勉強しながら作ることができました。

その作った詳細ー設計から作成の詳細についてなんですけれども、20世紀後半の新聞ということで、毎日新聞を対象としました。これには、通時的なコーパスを作るためには、一番古いCD、使える電子版のデータがどこにあるのかと調べたときに、毎日新聞が91年で一番早かったということもありまして、毎日新聞を選びました。そこから10年刻みで、1950年、戦後間もない頃から10年刻みで60年から2000年まで作ったというのが「毎日新聞経年コーパス」の第1版になります。基本的には、連載小説ですとか広告ですとか、そういったものを除いた記事のみとなっております。91年と先ほど申し上げたんですけれ

ども、この1年の誤差はちょっと自分の中ではもやもやが残ってはいましたけれども、やはり大規模なものを作りたいということで、こちらはぐっと我慢して、91年、2000年というふうにCDの方を利用させていただきました。

これは、第1版としては、自分の中ではかなり大規模なものではあったんですけども、1点問題がございまして、1950年というのは戦後間もない年ですので、そもそも新聞のページ数が非常に少なく、基本的には毎日2ページしか新聞が作られていませんで、祝日で4ページとなっています。なので、例えば2000年のものと比べると5倍近く数が少ないということがありまして、そこをそのまま使うということはちょっと難しかったので、ちょっとずつ増補していたということになります。2009年の第2版が毎月3日分に増やして、その後、2010年分を追加していて、一番少なかった50年を徐々に増やしていったというようになっております。

コーパスのサイズなんですけれども、こちらの表1のとおりです。現在、最新版としては第4版となっております。50年、60年はまだちょっと足りないんですけども、全体的には1,300万語を超える大規模なものとなっております。なかなか作るのは大変だったんですけども、まだ現段階で経年コーパス、このぐらいの時代で、経年で作っているものは、管見の限り、これが一番大きいサイズになっているのではないかと考えております。現在も、いろいろ増補しながら、中身も見ている最中でございます。

では、3点目、結局、作ったコーパスから何が分かったのかというところに移ります。

1) 番、先ほど雑誌の2種について申し上げたんですけども、新聞についても同じく調査した結果、新聞においても外来語が基本語彙の中に進出しているということが確認できました。ただ、こちらの傾きが—その速度というのは、雑誌に比べると若干遅いようです。ちょっと見やすいかなと思って白黒をカラーにしておりますけれども、お手元の資料の内容は同じです。こちらの青の部分の上位2,000語がもう少し上のところに来ていて、下の方が上位1,000語になっているんですけども、増えているということが確認できております。

2点目、基本語彙に進出している外来語の中身は結局どうなっているのかということなんですけれども、具体名詞と抽象名詞と比べると、抽象名詞の方が多いということが分かりました。具体名詞か抽象名詞かの分類—分け方につきましては、国立国語研究所で出している分類語彙表—単語を品詞とともに意味別に分類しているシソーラスと言われるものです—それに番号を付与しているんですけども、1.4が一番定型的なもので、1.2、

1.4、1.5を具体名詞とし、1.1、1.3を抽象名詞としたときに、抽象名詞の方が多
いということが明らかになりました。

次、3点目です。基本語彙に進出した外来語とその類義語との量的な関係に様々なものがあるということが分かりました。例えば図3の「テーマ」というものですが、こちらです。「テーマ」はオレンジ色になります。そのほかは外来語のこちらの「テーマ」の類義語に当たるものになります。オレンジ色とその他でまとめて見てみますと、この「テーマ」という外来語は、外来語が「主題」ですとか「題材」ですとか「題目」、「話題」といったほかの類義語に取って代わるような変化を見せています。ここですね。ここが収束されていくというような形になっています。といったものもあれば、例えば図4の「ケース」です。こちらになりますけれども、こちらは外来語が一部の類義語と並び立つような変化を見せているということです。ここですね。「ケース」が、このオレンジ色が「事例」ですとか「場合」ですとか「例」と、特にこちらの青い色の「例」と並ぶような感じで、一方で、減るものも一応ありますけれども、こちらの緑の「場合」、これはちょっと、「場合」は減っているのという疑問を抱いていらっしゃるかなと思うんですけど、これについては後ほど触れます。あと、こちらの黄色が「事例」になりますけど、これは数が少ない方で変化が起きているような状態です。というのが3点目。

4点目。一方で、多義語化していて、ほかの類義語の上位語になるような形で基本語化する外来語が一定数あることが分かりました。例えば「トラブル」という外来語があると思うんですけど、「トラブル」の意味は、辞書引くと、実はこれが面白くて、「トラブル」という辞書の見出し語のところを引くと、「事件」、それから「不祥事」、「けんか」と書いてあるんです。その語を探していくと、「トラブル」というので、トートロジーというんでしょうか、循環しているということが問題でして、自分でコーパスの中にある用例を分類してみました。そしたら、次のような3種6類に分けることができました。例えば、1番、人と人のトラブルです。例えば、①のように、誰と誰がけんかをしてみたいな、いざこざになるというようなことですね。実際ぶつかってみたいな、そういった物理的な、身体的な接触があるような場面で使われるもの。「出来事のトラブル」と名前を付けています。そうしたのもあれば、実際、表面的には、けんかしたというようなことがちょっと前後あったりする場合もあるんですけど、抱えている場合—何か悩みとして抱えていたとか。例えば交友関係でとか、男女関係でとか、〇〇関係でトラブルを抱えていたみたいなの、「関係のトラブル」というのがありました。2番目としては、モノのトラブルです。モノというのは、例えば機械が典型的

なものなんですけれども、エンジンにトラブルが起きたとか、どこどこにトラブルが起きたというのがモノである場合です。それがモノ、体も一種の機械と捉えられると思うんですけども、それが体になる場合もモノとして、「モノのトラブル」として分けました。例えば肌のトラブルですね。皮膚トラブルとかといったものになります。3番目はモノゴトのトラブルです。こちらは、英語にはなかなかないと言われるようなものがあつたりしますけれども、例えば、こちらですね。試験とか、新しい何かシステムとか法律とか、走り出したところで何か問題が発生した。ある運営や運用がうまくいかないような、不祥事が出たとかというようなものです。それが事故や事件だったりもします。それが⑥になります。線路内に乗客が下りるトラブルといったようなものになります。

これを通時的に調べてみたら、これらの意味は当然、全てが最初からそろっていたわけではなくて、次の表2のように、空欄を、空白を埋めていくような、つまり広がっていくような形で現れるということが分かりました。この埋めていく、広がっていくプラス量も増えているということが分かりました。例えば人と人のトラブル、出来事のトラブルが60年に1個だけ現れる。ちなみに、50年は使用例自体がゼロということです。なので、少なくとも経年コーパスの中では使用例がなくて、60年から増えてきたということです。60年に一つ出来事のトラブル、AとBがけんかしたというようなものがありまして、②の関係の方に広がって行って、機械が現れて、運営・運用が少しあるというような形です。80年になって、④の体の、肌のトラブルが現れてという感じです。埋まって行って、数も増えていくというような様相を見せています。ちょっとこちらの隣の方にそれぞれの意味・用法の類義語を付けております。例えば「けんか」、「口論」、この辺りが①になります。「もつれ」とか「不和」とか「不仲」とかが2番になります。「故障」、「不具合」もここに入りますね。「不調」ですね。「不調」は身体も同じですけども、「疾患」。これはスポーツ面で、スポーツ選手が肩の故障とかといったときの「故障」です。5番、6番は「障害」ですね。「支障」、この辺り。「不祥事」、「事故」、「事態」もあります。

このように、多義語化によって「トラブル」というのが、こちらの方の表2にあります3種6類に共通する意味として、抽象的な意味としてまとめられるのが、例えば「深刻・決定的な危機的事態に至る可能性を持って」ということですね。つまり、一步間違えればすごいことにつながりそうなものが「顕在化した不正常的な事態」というような共通する意味を導くことができまして、こういった抽象的な意味を表す基本語として外来語が多義語化していくんです。そうすることで、右の方の類義語、こちらの方を使うような場面で、その上位語の

ところの位置に立つようになって基本語化していくというような感じになります。これは類義語の方をちょっと御覧になると分かるかと思うんですけども、細かいところで述べているわけですね。つまり、類義語が分担して表していた、より具体的で限定的な事態ですね。要するに、悪いこと。新聞では悪いことをよく述べるわけですから、そういった悪い方の限定的な事態というところで、全て「トラブル」の一つの語で表すことができるということが恐らく原因として挙げられるかと思います。

また、4)番のようなタイプとはまた別に、別のタイプとして、特定の用法で類義語を大きく上回ることによって基本語化する外来語もあるということが明らかになりました。例えば「ケース」という語なんですけれども、「ケース」は、いわゆる入れ物の意味ではなくて、例えば「事例」ですとか「例」とか、そちらの方の意味になりますけれども、この抽象的な意味を表す外来語の「ケース」というのは、特定の用法にだんだん用法が集中してきて、意味役割分担をほかの類義語としていくというようなパターンです。その前にちょっと、「ケース」に今現在どのような用法があるのかを少しまとめておきます。AからDのような四つに分けることができるんですけれども、例えば単独の用法です。つまり、修飾部分がないようなものです。「ケースによっては」というものです。Bが、合成語の構成要素として、合成語の中の一部ということです。なので、例えば「特殊ケース」、「例外的ケース」とか「裁判ケース」とか、「〇〇ケース」といったものです。「ケース〇〇」は例がほとんどありません。「〇〇ケース」-後ろの部分の方で使われるものです。三つ目が、名詞句の被修飾語として使われる場合です。「仙台地裁のケース」、例ですね。それから、「悪質なケース」というようなものが三つ目です。四つ目は、修飾される前の部分が節である場合です。連体修飾節構造という言い方をするんですけれども、その中で使われる場合です。なので、例えばこちらですと、このケースの中身が前の部分に来るということです。例えば「極端な食事制限のストレスから、拒食症や過食症といった摂食障害を引き起こすケース」、ケースの内容に当たる部分になります。

そして、これらの用法をコーパスで、ちょっとその推移を見てみたいと思います。表3です。ね。「ケース」も「トラブル」と同じように50年は使用例がありませんでした。こちらの60年の方で単独用法のAからDまで一応全部そろっていました。数がちょっと凸凹ではあるんですけれども、こちらのDの方が、数が増えてきているということが確認できます。Cもちょっと凸凹ではあるんですけれども、まず、これが一つです。

次は、ほかの類義語とどのように使い分けているのかというのを調べてみました。そうし

ますと、この図5-1が単独用法なんですけれども、こちらは下の青い部分のこちらの「ケース」が増えてきているということです。「例」も一応ありますね。一応というか、部分があります。「場合」は、単独用法で、「〇〇の場合」だけというのは例がほとんどないわけなんですけれども、この二つが顕著です。

次は合成語の中身－構成要素となっているものです。こちらはオレンジ色の「例」がやはり多いです。裁判、どここの例、何とか例、前の部分が〇〇したケースというので例、あるいは何とかケースみたいなものがここですね。青とオレンジになります。名詞句の被修飾語である場合はこちらになります。これ、「例」が多いですね。これ、「ケース」になります。一番大事なDの部分になります。これは、この割合を見ますと、やはり「ケース」が増えてきているということが分かります。これは減りませんけれども、「例」も、〇〇する例が多いとかあるとかといった形で使われるものです。「場合」のことなんですけれども、「場合」は、合成語の構成要素としてというのは使いづらいわけですし、助数詞のように、例えば「3,000例」とかといったものも「場合」はちょっと使いにくいということもあったりして、いろいろ、かなり制限があるということが分かっています。このように、「ケース」は連体修飾節構造の被修飾語(同格連体名詞)、先ほど申し上げたように、内容を表しているような用法を大きく増やしていくことで基本語化したことがまず考えられます。その「ケース」の意味の部分なんですけれども、次のような例のように、「(すでに起こった)良くないコトガラ」、つまりケースですね。その事態が多いとかあるといった型を獲得してきていることが分かりました。例をお見せした方が分かりやすいかなと思うんですけれども、例えば「児童相談所が事態を把握しながら救えなかったケースが、98年度は判明しただけで8件あった」です。これはやはり新聞で述べるような、よく使われるようなパターンを獲得しているということです。「積極的な対応をちゅうちょしたり関係機関との連携がうまくいかないケースが依然多い」といったものになります。それに対して類義語の「事例」ですと、例えば「活用事例」のような結合用法が多い。それから、「例」というのは助数詞の用法が顕著です。「300例」とかです。数を数えるときに使われる。それから、「感染例」のように、これも「事例」と同じく結合用法が顕著。それに対して、「場合」という類義語は提題用法によく使われているということが分かりました。提題用法というのは、トピックを提示するときのものとして、例えば「今回の場合」、次に「、」を入れるやつですね。「彼の場合」、トピックを提示するときの用法に顕著です。それから、もっと顕著なのが、仮定条件的な用法にかなり用例が使われていました。例えば「次のようなことが起きた場合には」といったところに力を発

揮っていて、つまり、すみ分けているということです。ある用法ではこちらの語、ある用法ではこちらの外来語みたいなどころで、役割分担をしながらすみ分けていくということが明らかになりました。そこで、「ケース」というのが形式名詞化していくわけですね。

次、6番目。次は、文章構成機能を担う外来語が増えているということが明らかになりました。こちらはテキスト構成語と従来言われるものとして、例えば、日本語の中に指示詞の「この」というものがあるんですけども、この指示詞である「この」がどのように使われるかと言いますと、前で述べていたものを、結束性と言うんでしょうか、引き受けて、次の表現を次に述べていくときに、前の部分を引っ張って関係付けながら後ろのところまで述べていくというようなときに「この」という指示詞がよく使われるんですけども、つまり前方照応ですね。文章内での前方語句の再表現というものに「この」というものが使われるんですが、例えば、1番の「この」というのが、ここですね。エンジンというものが、ここですね、「予備の燃料を持ったエンジン」としたときに、「このエンジンによって」と読むんです。だから、これはただ繰り返しているだけなんですけれども、次に、前の表現を別の語句に言い換えるパターンです。例えば、「うとうしい梅雨期、心身ともにゲンナリする季節だ」としたときに、この「季節」を「シーズン」と換えるんです。「しのぎにくいこのシーズン」というふうに言い換えるパターンです。三つ目が、前で述べたものを後ろの叙述につなげていくときに捉え直し、つまり、書き手がちょっと別の表現を、解釈したりとかして捉え直すパターンです。文章の前方で述べられている叙事(事柄)をコノ語句で名詞化しながら後ろにつなげていくというようなものです。例えば、「ローラー・ゲームがテレビで放送されたせいで、去年は男の子の間にローラー・スケートの売れ行きが好調だった」、「このブーム」というふうに捉え直すというようなものです。

この前方照応しながら後ろの部分につなげていくような、ここでは「コノ語句」という言い方をしたんですけども、「コノ語句」がどのように変わっていくのかというのを外来語に対して調査してみた結果です。その結果が表4となります。第2版を使って、単純語のみを調査しました。その結果、こちらになります。実数が表4のとおりで、それを100万字当たりの出現率として換算したものが次の図6となります。ここですね。下の方がここですね。時代を下るにつれて、こちらの方の捉え直しの青の部分、1950年はやはり数が少ないということもありますし、若干別な感じなので、ちょっとこれは除いて、60年から見ると、ここが広がってくるということが分かります。つまり、文章構成機能を担うような部分に外来語がどんどん使われるようになったということです。これは、例えば繰り返しですとか言い換

えの場合は別に具体名詞でも抽象名詞でもどちらでも使うわけです。「このエンジン」とも言えるし、「このシステム」としたときに、どちらでも使えるんですけども、3番の捉え直しというのは、語より大きな単位で表現される事柄がその直前にあるわけですよ。それを名詞化して1語で表さないといけないわけですから、基本的には抽象名詞でなければなりません。なので、この結果ですね。「コノ語句」による再表現のこの方式を調べたこの結果は、結果としては、3番の捉え直して文章構成機能を担うことができるものとして抽象的な外来語が増えてきた、あるいは増えてきたことが言えるようなことがあると思います。従来で言うと、このような機能を担うものとして、高崎みどり氏は、これを「テキスト構成語」という用語を使うんですけども、正にそこに外来語がどんどん増えてきているということが明らかになりました。

お時間ちょっと迫ってきたかと思えますけれども、4番。では、なぜコーパスが必要なのかということ。3節で、1)から3)番は別に語彙表だけでもできる研究です。それに対して4)番から6)番は、原文を入力した、かなり数もある、規模のあるものじゃないとできない研究です。このうち、「トラブル」に見られる多義語ですね。3種6類で上位語化していく、その類義語の上に立つようなタイプと、「ケース」に見られるようなすみ分けタイプがありますね。これは実は、その危険性というか、現在そういった危ないことが起きているよということを指摘している研究があります。樺島忠夫氏によると、こういった危ないことが起きていて、何か訳の分からないことが起きているよという指摘です。例えば、既存の適切な語によって表現される労力を省いて、覚えた外来語、例えば「アクセス」を乱用すると、その外来語の使用が一般に広がるにつれて、次第に多義語化して、「接続」、「利用」、「交通手段」のような、文脈上で適切な意味の区別をぼやかしてしまう。それから、例えば、「開く」、「削る」、「切る」といったものも、「カット」で全部言えてしまうと。これ、危ないんじゃないかというようなことです。こうしていくと大和言葉も消えていくよというような趣旨です。その危険性について警告しているものということですね。

次は、私のコーパスの調査から明らかのように、実は新聞には、既存の類義語が当然あるわけです。それまでに従来の類義語があるわけなんですけれども、それにもかかわらず基本語化している抽象的な外来語がたくさんあるということが明らかになりましたけれども、では、新聞に、なぜ必要がない外来語が増えていて、それを基本語化している必然性があるだろうかという問題になります。

それには、例えば「トラブル」なんですけれども、その多義語化によって——ここですね。

共通する意味として、深刻・決定的な危機的事態に至る可能性を持って顕在化した不正常的な事態—ちょっと長いんですけども、要するに、新聞にたくさん取り上げられるようなものに、こういったものがあるわけですね。こういった抽象的な意味を獲得して、事態は、詳しい内容はまだ分からないけれども、でも報道しないといけないという文脈で、「(何らかの)トラブル」と言えるわけですね。つまり、細かいことが、それがけんかだったのか、結局事故だったのか事件だったのかまだ分からない段階でも、まだ言える。何らかの何かが起きた、不正常的な事態が起きたということを新聞で報道しないといけない。そのときに便利に使える単語であった可能性があるというわけです。

一方で、「ケース」というのもあるんですけども、似たような感じですね。「(すでに起こった)良くないコトガラ」、ケースと言えものが、実際あるとかないとか、起きたとか想定できるといったところは、新聞でたくさん多用される表現で、もう型として獲得されていくというようなことです。

とすると、「トラブル」も外来語も新聞にとっては、「使う必要のない単語」とは、とてもじゃないですけども、言えません。新聞で基本語化している抽象的な外来語というのは、新聞の文章構成において「必要」である単語であるということですね。ただ、ほかにもいろいろあると思うんですけども、それを実証するためには、直感とか、ちょっと調べてみたというのはやはりできませんで、量質ともにちゃんとした良質の経年コーパスが必要になってくるということです。そうすることによって、計量的な語彙史研究もできるということが言えるかと思います。

最後に、今BCCWJの拡張事業が進行中だと伺っているんですけども、そのコーパスが大規模になることで、私のように、やりたい研究ですとか、あるいは明らかにしないといけないような研究というのが、良質のコーパスがあることによって、それができるわけですね。そういったことで、量質ともにそういった研究が増えていくことをとても期待しています。また、作ってくださる方々に是非お礼を申し上げたいと思っています。

これで発表は終わりにしたいと思います。拙い発表で大変失礼いたしました。ありがとうございます。

○相澤主査

金先生、興味深い御発表、大変ありがとうございました。

では、ただ今の御発表につきまして、御質問がございましたら、よろしく願いいたします。

す。オンライン参加の方も挙手にて御質問いただけますので、よろしく申し上げます。いかがでしょうか。発表自体についての御質問はよろしいですか。

○前川委員

じゃあ、お願いします。よろしいですか。

○相澤主査

はい、お願いいたします。

○金発表者

はい、お願いします。

○前川委員

国語研の前川です。お久しぶりです。

○金発表者

御無沙汰しております。ありがとうございます。

○前川委員

面白い発表を聞かせていただいたと思うんですけど、この発表の中での「基本語」というのは、基本的には高頻度語のことだと考えていいんですかね。

○金発表者

そこはちょっとやはり厳密に最初に申し上げるべきだったんですけども、高頻度で幅が広い、レンジが広い、つまり、話題を選ばないという意味で御理解いただければと思います。

○前川委員

高頻度は調べれば、ある意味簡単に分かるんだけど、その「幅が広い」という辺りの判断は何か客観的にできるものなんですか。

○金発表者

話題の分け方というのは非常に、自分では判断できませんので、新聞を選んだ理由でもあるんですけども、紙面を一つの話題、レンジと捉えました。

○前川委員

なるほど。なるほど。

○金発表者

そうすると、いろんな紙面で使われるという意味です。

○前川委員

そうすると、何十年か見ていくと、その基本語に出たり入ったりがあるような気がするんですよね。一時的に入って出ていくとか、外来語の場合、そもそも消えてしまうのもあるかもしれないけれども、そういったののリストがあると、私、ちょっと個人的に見てみたいなという気がするんですが、それは参考文献の愛蘭さんの御本の中に、見ると、出てくるということでしょうか。

○金発表者

去年出した本の中にはちょっとそのリストは載っていませんけれども、博論が現在公開中で、凡例のところで、減ってきている語の御確認はしていただけます。

○前川委員

分かりました。拝見します。

○金発表者

やはり、どんどん力を獲得するものと、どんどんしなびていく語と、両方あります。

○前川委員

ありますよね。

○金発表者

はい。ありがとうございます。

○前川委員

どうもありがとうございました。

○金発表者

ありがとうございました。

○相澤主査

ありがとうございます。

ほかにもございましたら、よろしく願いいたします。よろしいでしょうか。

(→ 挙手なし)

では、改めて、御講演を頂きまして、どうもありがとうございました。

○金発表者

ありがとうございます。

○相澤主査

ありがとうございました。

それでは、残りの時間、20分ほどございますが、意見交換に移りたいと思います。本日お二人の先生方に御発表いただきましたけれども、それに基づきまして、あるいはそれを離れましても、それぞれの委員の皆様の御関心に基づきまして、言語資源やコーパス等の活用という観点も踏まえて御意見を頂ければと思います。

○成川委員

いいですか。

○相澤主査

はい。お願いいたします。

○成川委員

共同通信の成川と言います。

先ほどの「ケース」、記事を書いている立場から言うと、どういうときに使うかというのをちょっと説明した方がいいかなと思いました。「例」ですとか「事例」とか「場合」とかと続くのが嫌なので、「場合」の後は「ケース」を使う、「ケース」を持ってきたら、「ケース」、「ケース」、「ケース」とは行かず、「ケース」じゃなくて「場合」にする、「事例」にするという、選択肢の一つとして非常に有用ですね。書いている立場から、どういうとき使うかと言うと、同じ言葉が続けないために使っています。

○金発表者

ありがとうございます。現場のお声が聞けてとてもうれしいです。やはり新聞の中で、結束性を持ちながら、同じ語を避けるというのは、これはとてもいい文章の一つだと言われていまして、そこにやはり外来語という選択肢が更に加わって多様な表現ができるというところでも、先ほど申し上げた「コノ語句」のところでも似通うところがあるかなと思いました。ありがとうございます。

○相澤主査

ありがとうございます。

○石川副主査

一つよろしいですか。

○相澤主査

はい。お願いいたします。

○石川副主査

石川です。

言語資源整備という観点から、今日二つの御発表を伺って感じたことを申すと、外来語

の問題を科学的に議論していくには、やはり、一般公開された経年コーパスが必須であるということです。金先生は大変なお手間を掛けて、御自身の研究用に経年的な新聞データベースをお作りになったわけですが、同様のものが幅広い研究者に使えるようになっていくことが重要だと思います。現在ですと、朝日新聞や読売新聞などにオンラインの検索サービスがありますが、全体をダウンロードして形態素解析するというのは権利的にも技術的にも困難です。この点に関して、各新聞社が持っている自社の記事データを言語学的にアノテーションし、共通のプラットフォームに載せて、言語研究に使えるようにするといったイニシアチブが出てくると、外来語のみならず、日本語研究の更なる発展に有用であるのではないかと感じた次第です。

○金発表者

ありがとうございます。

○相澤主査

ありがとうございます。

よろしいでしょうか。

では、本小委員会恒例となっておりますが、最後に皆様に一言ずつ順番に、本日の感想等を踏まえてお伺いしてまいりたいと思います。石川副主査はただ今御発言されましたので。

○石川副主査

はい。ありがとうございます。私のコメントは既に申し上げました。

○相澤主査

では、言語小委員会から、あいうえお順とさせていただきます。本日、ローマ字小委員会の委員の皆様にも、お時間の許す限りお願いしたいと思っております。では、神永委員、よろしく申し上げます。

○神永委員

神永です。小椋先生、金先生、とても面白い、興味深いお話、ありがとうございました。

まず、小椋先生のお話を伺いまして、これ、正に辞書に関わる大きな問題ですので大変興味深く拝聴した次第なのですが、なぜ辞書に関わるかと言いますと、辞書の見出し語の語形というのは、こうした調査を参考にして決定することになっているからなんです。小学館の『現代国語例解辞典』の事例というんでしょうか、ケースというんでしょうか、御紹介いただきましたけれども、この辞典もBCCWJでの調査結果を基にして外来語の見出し語の語形を決定しています。外来語の表記の場合、辞書では、見出し語の語形を決めるだけではなくて、表記の異なる異形をどう扱うかという問題もとても大きな問題としてあります。デジタル辞書の場合には、こうした異形からでもできるだけ引けるようにデータを埋め込んでいるんですけども、紙の辞書の場合にはそうはいかないわけなんです。異形と考えられる語形をどこまで載せるかということも、こういった調査、研究からまた決められるのではないかと考えています。紙の辞書の場合、こうした異形と言われているもの全てを載せるわけにはいきませんから、やはりある程度標準的な表記は決めつつも、揺れとして比較的広く使われているものをコーパスなどから拾い上げて、柔軟に対応できるようにしておくべきだろうと考えています。おっしゃっておられたように、コーパスの継続的な拡張による言語の使用実態の把握が必要であろうと思います。そして、それを辞書の中で定期的にと言いますか、頻繁に反映させていただくような形にできたらいいかなと思っています。

あとちょっと、私の話、これ長くなってもいけませんので、あれなんですけれども、もう一つちょっと大きな問題がございまして…。長音表記にするか連母音表記にするかということなのですが、これ、辞書の配列に関わってくるんです。「メイン」か「メーン」かに関わってきますので、これも大きな問題なので、どちらが今一般的なのか、あるいは、もう両形あるんだったら、それはそれなりの対応が必要なのかということをちょっと考えさせていただきたいなと思いました。

それから、金先生のお話なんですけれども、これも大変興味深く拝聴したんですが、特に「ケース」の問題。実はこれ、『使い方の分かる類語例解辞典』というのをちょっと御紹介いただきましたけれども、これ、個人的なことなんです、私が編集を担当した辞典でございまして、これ、元々1993年に出した辞書なんです。金先生がお使いになった辞書は2006年の新装版なんです、内容は全く変わっていません。それで、お示しくださった図4を見ますと、ちょうど1970年ぐらいから「ケース」の類義語の量的推移がちょっと上がっているというようなことをお示しになってはいますけれども、正にこういったところがありまして、類語例解辞典の中では「ケース」と「事例」と「例」というのを比較してみたんです。こ

ういったように、外来語が既存の類義語の一部の用法への進出というようなことがあるように思われますので、これもまた、辞書の中では意外と外来語というのは扱いが今、簡単と言いますか、本当に対応する日本語の意味の置き換えだけにしてしまっていますので、そのところをもう少し考えなければいけないのかなというようなことを考えました。

以上、辞書編集者としての感想です。どうもありがとうございました。

○相澤主査

ありがとうございました。

では、言語資源小委員会でオンラインで御出席の方からということで、武田委員、よろしくをお願いします。

○武田委員

武田でございます。本日は小椋先生、金先生、大変貴重な御報告、ありがとうございました。

小椋先生、金先生の御報告ですけれども、前回の石川副主査の御報告にもありました論点、言語資源と外来語、それから言語資源の整備にも密接に関係するテーマであったと思います。また、今後の議論、検討の方向性にも資する内容であったと考えております。是非、幅広い見地で言語資源の活用ということを検討していただくことができればと考えております。

以上です。

○相澤主査

ありがとうございます。

では、オンラインで御参加の齊藤(美)委員、いらっしゃいますか。

○齊藤(美)委員

はい。齊藤(美)です。

○相澤主査

お願いします。

○齊藤(美)委員

小椋先生、金先生、御発表どうもありがとうございました。大変興味深く伺いました。

外来語について今後国語分科会で検討する可能性があると思われ、冒頭で相澤主査から伺ったかと思えます。使用や表記の指針、昨日告示された「ローマ字のつづり方」のようなものが外来語についても何かあると便利だろうなどは思います。日本語の書き言葉でのコミュニケーションあるいは日本語教育にとって重要になると思えます。日本語学習者にとって片仮名語は分かりづらいと聞くことがありますので。しかし、例えば、全ての語について表記方法を提案するというのは、数が多過ぎて無理だろうとも思われます。かといって、ルールを示すのも難しいだろうと。でも、それぞれの各語によって、こちらの表記が自然かな、なじみがあるみたいなのは、個人差はあると思うんですけど、あるのかなという気もしていて、今日の小椋先生の御発表にあった事例で言えば、同じ連母音でも、「メール」というのは私は違和感がちょっとあるんですけど、「プレイ」は比較的なじみやすい印象でした。そのルール化の難しさというのは、今日石川副主査が小椋先生に対してなされた御質問とその小椋先生のお答えの中でも触れられていたかと思えます。そうすると、外来語表記の統一、揺れをなくすということ、例えば目指すことは必要なのかとか、そういったことも考えることが重要になってくるのかなと思いました。

明治期の文献を読むことがあるんですけども、そうすると、言文一致が始まった頃ですから当然なんですけれども、外来語、片仮名語以外の漢字使用の有無とか送り仮名とか、いろいろなことが本当様々であって…。当然なんですけど、でも、特に、ちょっと読みづらい気もしますが、でも、これはこれでいいのかなというような気もしていて、大学で教員をしているので、学生の論文添削を現代でも例えばするとき、表記が揺れていると、統一してくださいと言いたくなるんですけど、でも、そこはそんなに重要なのかなという気もしたりもしていて、本当に必要なのかなとか、いろいろ私も迷うところがあって、これは、言葉はどうあるべきかという、それぞれの使用者、学習者も含めると、本当言語のレベルはいろいろだと、習熟度はいろいろだと思います。そのそれぞれの使用者の考え方と大きく関わることだと思いますので、今後、慎重な議論が必要そうだなと思っておりました。

以上です。ありがとうございました。

○相澤主査

ありがとうございました。

では、前川委員、よろしくお願いいたします。

○前川委員

小椋さんの発表を聞いていてちょっと思ったことというか、発表についてということじゃなくて、私が自分で、これどうなっているんだろうと思うことが一つあったので、それを述べてコメントに代えさせてもらおうと思うんですけど、外来語の表記というときに、よく原語との関係というのが、今日もお話の中で出てきたと思うんですけど、ケースによりますけど、長母音の場合、今日示されていた例はほとんどが英語の例だと思うんだけど、英語には実は長母音、短母音の対立はないんですよね。母音の明暗の対立はありますけれども、基本的に、「プール」と「プル」とか、「ピック」と「ピーク」とかそういう、我々はそれを長母音、短母音と聞いているんだけど、元々は音色の違い、長さも違うけど、むしろ音色の方が違うと、そういう対立を我々が勝手に長短と聞いて、それで混乱しているという状態だと思うんです。だから、ほか、そうでないケースももちろんたくさんあるわけだけど、長母音に関して言うと、基本は日本語の問題なのかなという気がちょっとしています。ほかにもきつこういう、対立を日本語に置き換えて聞いて、それが問題になっているということがあるんじゃないかなと、具体例挙げられませんが、そんなことを思いながら聞いておりました。

感想です。

○相澤主査

ありがとうございます。

では、森委員、おられますでしょうか。

○森委員

日本テレビ、森です。カメラオフのままで失礼します。

まず、小椋先生の発音と表記について、アナウンサーとしてとても興味深く伺いました。お話の中にもありましたけれども、新聞やテレビは各社、表記を統一しておられて、これまで長音表記だったものを連母音に換える流れにあるなということは感じております。ほかに、弊社で言いますと、「マネージャー」という言葉が、これまでは「マネジャー」だけしか

認められていなかったものが、今年「マネージャー」が認められて、さらに、「マネージャー」を優先して使うということになりました。実際、世の中でどういう言葉が使われているかというのを優先してそうなったというところでもあります。

石川副主査から少しあったんですけど、外来語ではなくて、日本語の「冷凍食品」というのも、表記は平仮名で書くと「れいとう」なんですけど、発音としては「れーとー」と引っ張る長音になっている。私は関西出身で、アナウンサーになったときにアクセントの違いや無声化や鼻濁音ということを徹底的に教わったんですけども、例えば「稽古」とか「警察」という言葉が、「けいぎつ」ではなく「けーぎつ」である、「けいこ」ではなく「けーこ」であるということも厳しく指導を受けました。共通語としては、エイの音ではなく、エーと引っ張る音で発音するんだと聞いたんですね。外来語は、表記が「メーン」で、ただ、実際の発音は連母音が増えてきている。逆に、日本語は、連母音表記なんだけれども、実際の発音は共通語では長音であるという逆の動きがあるなということをお今日の小椋先生の発表と石川副主査の質問から感じておりました。

それと、金先生の発表についてもすごく興味深く伺っていたんですけど、これはBCCWJ 2のお話の時もそうでしたが、新聞が資料に入れられなくなった、各社が有料の検索サービスを作っているためにそれが今できなくなっているということを伺うと、この金先生のようなコーパスの作り方も今後難しくなっていくんだろうか、それはとてももったいないことだなとやはり、今日もそのことに関しては感じました。

以上です。

○相澤主査

ありがとうございました。

では、ローマ字小委員会からはオンサイトで御出席の3名の委員の方々から一言ずつ頂戴できればと存じます。私の不手際でちょっと時間が迫ってしまい、申し訳ありません。森山主査からよろしく願いいたします。

○森山主査

今日はお二人の先生の本当に興味深いお話、ありがとうございました。いつも思いますのは、正書法としての問題という点で、先ほどの「冷凍」をどのように書くかということなんかはローマ字とも共通する問題でもあります。また、外来語の場合、同じ語であっても、例

えば「make」という同じ英語の単語から来ている、顔のときには「メイク」で、「メーカー」というときには大体長音符号で伸ばすとか、「take」でも、「テイクアウト」というのは大体街では「テイク」と書いてあるのに、「ターンテーキング」みたいなときには大体「ターンテーキング」というふうに伸ばすとか、そういう語ごとの慣用というのが非常に大きくて、その慣用に対する扱いというものと、それから原理的にどう考えるのかということ、本当に難しい問題。そのために、非常にこのコーパスというのの重要性というのも改めて認識させていただいた次第です。どうもありがとうございました。

○相澤主査

ありがとうございます。

では、成川委員、よろしく願いいたします。

○成川委員

私たちは用語集を作るんです。新聞でいうと、新聞協会で大本を作って、基本はそうなんですけれども、今、真っ二つに割れている部分があります。それは「ウイ／ウィ」「ウエ／ウェ」「ウオ／ウォ」。朝日と読売は後ろを小さく書きます。共同をはじめ、共同の加盟社ですね。日経、産経、毎日含めて、あと全国の地方紙は、昔ながら大きく書いています。どちらでもいいんですけど、本当に。どっちかという朝日と読売は、もう最近はその流れだから、そっちに変えたという感じですよ。あと、さっきのお話でいうと、新しく入ってきた頃は2語で、二重で書くだけではなくて、やはり変化している。今お話あったように、お化粧は「メイク」、うちは「メイク」のままですし、あと、「プラットホーム」は駅、システムを作るときの「プラットフォーム」とか、同じ英単語ですけども、新しい意味のときは昔はなかった「フォ」を使うとか、もうとにかくぐちゃぐちゃで、基本的には元の単語の表記で基準を作って、それに当てはめようとしているわけですけども、なかなかそうも行かないので、今後は一語一語考えないといけないのかなと。

あともう一個言うと、うちの場合は外来語の表記集と外国人名・地名とあるんです。外来語は、よく考えると、最後がyのときは「ー」にする、長音にすると言うんですけど、これは元は何語なんだろう。vは「ヴ」にするといったって、ドイツ語と英語じゃ全然違うわけですし、非常に大変だなと今日の話も聞いて感じました。

○相澤主査

ありがとうございます。

では、前田委員、よろしく願いいたします。

○前田委員

本日は小椋先生と金先生のお話を伺って、大変勉強になりました。小椋先生の外来語の表記の御分析と、それから、金先生は外来語の意味とか、あるいは文法的な機能とかテキスト上の機能などの問題について、新聞を経年的に追われていらしたという研究でしたけれども、どちらの研究においてもコーパスというものが非常に有効である—有効であるだけじゃなくて、説得的な、ある種の科学的な結論に到達するためのツールになっている—ということを非常に改めて強く認識しました。

私も日本語教育に—先ほど日本語教育の話も出たんですけど—若干関わっておりまして、どっちの表記がいいんですか、棒引きがある方がいいのか、ないのかというのは本当にいつも質問されたりすることがあります。実際、現実にはどっちも使われているわけなので、こちらが正しいということは言えないんですけども、ただ一方で、やはり、こっちの方が多いんだとか、あるいはだんだん増えているんだとか、そういうデータがあると、どっちを使ってもいいけれども、現実はこちらなんだということをお教えることによって、学習者が自分で選ぶことができると、そういう意味での説得力のある教育というものにつながると思いますので、そういうことができるようにするためにも、やはりコーパスというのが非常に有力なものになっているのである—ということを改めて認識、本日いたしました。どうもありがとうございました。

○相澤主査

ありがとうございました。

まだまだ皆様の御意見を伺いたいところではありますが、非常に残念ながらお時間となつてしまいましたので、まとめに入りたいと思います。

本日、外来語ということで、ローマ字小委員会の委員の皆様にも御参加いただきまして、本当にありがとうございました。いつもとはまた少し異なる雰囲気の中、非常に密な情報交換を行うことができまして、今後にも是非結び付けていきたいと思っております。

御講演いただきました小椋先生、金先生に改めて御礼を申し上げます。

○小椋発表者・金発表者

ありがとうございました。

○相澤主査

ありがとうございました。

本日、議事、ここまでといたします。

本日の言語資源小委員会は、これにて閉会いたします。ありがとうございました。