

第13回 国語分科会言語資源小委員会(Web開催)議事録

令和 8 年 6 月 2 日

15 時 00 分 ～ 15 時 50 分

文部科学省15階 15F特別会議室

〔出席者〕

(委員)相澤主査、石川副主査、植木、神永、日高、前川、森、森川各委員(計8名)

(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、森川各委員及び事務局は15F特別
会議室にて参加。

〔配布資料〕

- 1 文化審議会国語分科会言語資源小委員会委員名簿
- 2 文化審議会国語分科会言語資源小委員会の会議の公開について(案)
- 3 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)
- 4 ドイツ語コーパス(DeReCo)拡充の背景

〔参考資料〕

- 1 文化審議会国語分科会運営規則
- 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)
(文化審議会国語分科会 令和7年3月17日)

〔経過概要〕

- 1 事務局から定足数の確認があった。
- 2 今期1回目の言語資源小委員会の開催に当たり、山田国語課長から挨拶があった。
- 3 事務局から配布資料の確認が行われた。
- 4 参考資料1「文化審議会国語分科会運営規則」に基づき、委員の互選により、

相澤委員が言語資源小委員会主査に選出され、相澤主査により、石川委員が副主査に指名された。

- 5 事務局から配布資料2「文化審議会国語分科会言語資源小委員会の会議の公開について(案)」の説明があり、言語資源小委員会の会議を公開することが了承され、以降の議事が公開された。
- 6 相澤主査及び石川副主査から、就任に当たっての挨拶があった。
- 7 事務局から配布資料3「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)」について説明があり、説明に対する質問及び意見を尋ねたが、特に発言はなかった。
- 8 事務局から配布資料4「ドイツ語コーパス(DeReCo)拡充の背景」について説明があり、説明に対する質疑応答が行われた。
- 9 配布資料1「文化審議会国語分科会言語資源小委員会委員名簿」の順に、自己紹介を兼ねた挨拶があった。
- 10 次回の言語資源小委員会について、令和8年6月30日(火)15時から17時まで、対面及びオンラインで開催する予定であることが確認された。
- 11 質疑応答及び自己紹介における各委員の発言等は次のとおりである。

○鈴木国語調査官

定刻になりましたので、ただ今から、第13回、今期1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。本日は最初の小委員会でございますので、後ほど主査を選出していただくまでの間、事務局で進行いたしますので、御了承のほどお願いいたします。

会議は、実際にお集まりいただく形でも、オンラインで参加していただく形でも可能な形式で今後も継続していく予定でございます。また、後ほど会議の公開についての確認が取れましたら、会議はオンラインでの公開となります。傍聴の皆さんがオンラインでこの会議を御覧になるということを前もって御承知おきくださいませ。

まず、定足数の確認を行いたいと思います。言語資源小委員会は9名で構成されており、本日は8名の御出席を頂いております。文化審議会国語分科会運営規則の第4条に定められている会議開催に必要な過半数出席という条件を満たしていることを御報告い

たします。

初めに、国語課長の山田から御挨拶申し上げます。

○山田国語課長

文化庁国語課長をしております山田でございます。今期第1回目の言語資源小委員会の開催に当たりまして、一言御挨拶を申し上げます。

言語資源の在り方につきましては、これまで令和5年9月の国語分科会で、「日本語のデジタル言語資源の整備に関する国語分科会の見解」という報告、そして令和7年3月の国語分科会で、「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」という報告をおまとめいただきまして、これらが現在文化庁で実施しております事業「信頼できる言語資源としての現代日本語の保存・活用のためのデジタル基盤整備事業」につながっております。こちらの事業では、書籍データを中心とした国立国語研究所の「現代日本語書き言葉均衡コーパス」ー略してBCCWJーこちらの拡充を進めていただいているところでございます。

現在、社会で様々なAI、そしてその基となる大規模言語モデルというものが幾つも開発されております。けれども、そういったものの大半はウェブ上の言葉をベースに開発されたということでございまして、現在、国語研で拡充を進めていただいているBCCWJ2とは本質的な違いがあると認識をしております。また、ちょうどこの土曜日にBCCWJ2の一部公開を記念したシンポジウムというのを国語研で開催していただきまして、私も参加をさせていただきました。この場で様々なお立場の方の御発表をお聞きしまして、BCCWJ拡充の重要性、そしてまた、様々な分野での活用の可能性ということについて、認識を新たにしたところでございます。

こうした中、今期も言語資源小委員会が引き続き設置されております。今期におきましても、国内での様々なコーパスに関する取組について取り上げていくということとともに、少し話題の幅を広げての御議論も期待をしているところでございます。委員の皆様におかれましては、大変御多忙の中とは存じますけれども、この小委員会におきまして、デジタル言語資源の整備と活用の現状、さらにはそこに現れた様々な課題につきましても御議論いただきまして、今後の国語施策に資するような御助言、御提案を頂けますと幸いです。今期もどうぞよろしくお願いいたします。

○鈴木国語調査官

それでは、議事に入る前に、本日の配布資料を確認いたします。

配布資料1「文化審議会国語分科会言語資源小委員会委員名簿」、配布資料2「文化審議会国語分科会言語資源小委員会の会議の公開について(案)」、配布資料3「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)」、配布資料4「ドイツ語コーパス(DeReKo)拡充の背景」、参考資料1「文化審議会国語分科会運営規則」、参考資料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)」、以上です。

(互選により、相澤委員が主査に選出され、相澤主査から副主査として石川委員が指名された。)

○鈴木国語調査官

ここからの議事につきましては、相澤主査にお願いいたします。

○相澤主査

それでは、配布資料2「文化審議会国語分科会言語資源小委員会の会議の公開について(案)」を御覧いただければと存じます。こちらは、4月30日の国語分科会で御確認いただいた「文化審議会国語分科会の会議の公開について」に倣った案で、言語資源小委員会の会議について、公開の在り方を定めようとするものであります。事務局から御説明をよろしくお願いいたします。

○鈴木国語調査官

配布資料2、会議の公開についての案でございます。基本的に言語資源小委員会の会議は公開するというので、その中身について定めようというものでございます。

まず、公開をしない事柄としては、今の主査・副主査の選任等の人事に係る案件、それから主査が、公開することにより審議に支障を及ぼす、そういう可能性を認める正当な理由があると認められる案件、この二つにつきましては公開をしないという形で、それ以外は全て、会議は公開となります。

会議を公開する場合は、その1週間前に文化庁のホームページにおいて情報を公開するという形になっております。

会議の傍聴につきましては、オンラインでの傍聴という形になりまして、文化庁国語課の定める手順によって登録をお願いする形を取っております。なお、登録の際に申出のあった場合、特段の問題がない限りは、録音、録画等も認めるということで対応いたしたいと思っております。もしも、何か会議の進行を妨げる行為等があった場合については、その会議の傍聴からの退場を命ずる等の措置を取ることができるという定めもございます。

会議の資料につきましては、原則として公開をするという形でございます、特段理由があるものについてのみ非公開ということが行われるということでございます。

議事録の公開でございますが、基本的に議事録は、この会議におきまして皆様のご了承を得られた後、公開するという形でございますので、皆様の目が通った上での公開になるということになります。一部非公開とする部分が生じる場合、その部分については議事要旨を作成して、議事録とは別に公開をして対応することが最後の項目として挙げられております。

以上でございます。

○相澤主査

ただ今の説明について、御質問等ありましたら、お願いいたします。

(→ 挙手なし。)

よろしいでしょうか。では、御異議ございませんようですので、文化審議会国語分科会言語資源小委員会の会議の公開については、合意いただいたものといたします。

これに従いまして、ここからの議事については、オンライン上で公開することにいたします。

(ユーチューブ配信開始)

では、配信開始ということで、ここから傍聴の方々も会議の様子を御覧になられるということです。

言語資源小委員会の主査を拝命いたしました相澤から一言御挨拶を申し上げます。

先日、BCCWJ2の公開を記念したシンポジウムに参加をする機会を頂きました。国立国語研究所からBCCWJ2の基本設計等を御紹介いただいた後に、ゲストスピーカーの先生方から、例えば表記の変遷ですとか、日本語におけるコロケーションのハンドブックについてですとか、英語圏の均衡コーパスの歴史や、それを踏まえた日本語言語資源、BCCWJ2プロジェクトの意義等をお話いただきまして、非常に興味深く拝聴いたしました。

本小委員会でも、まだ議論すべきことが多々あると感じております。今期もどうぞ、皆様、よろしくお願いいたします。

続きまして、副主査である石川委員の御挨拶を頂きたく、よろしくお願いいたします。

○石川副主査

副主査に任命を賜りました神戸大学の石川でございます。皆さん、はじめまして。

本小委員会でのこれまでの審議を通して、言語資源が非常に多様なもので、その有用性も多岐にわたることが見えてきたところでございます。今後については、BCCWJ2のような核になる大規模言語資源をしっかりと整備していくことと並行して、それぞれにユニークな特性を持った各種の小規模コーパスにも目配りをしていくこと、また、それらの研究成果を統合して社会ニーズの解決にうまくつなげていくこと、こうした点を皆様とともに今期の小委員会において深く議論できればと存じております。

以上です。

○相澤主査

ありがとうございます。

まずは、言語資源小委員会から出しております報告のフォローアップということで、前期の審議を進めてきた言語資源小委員会の活動を踏まえまして、今期の主な審議事項について確認をしたいと思います。その後に具体的な審議に入っていくことといたします。

配布資料3「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)」についての説明を、山田課長からよろしくお願いいたします。

○山田国語課長

国語課長の山田でございます。それでは、今期の本小委員会での審議事項を御説明いたします。

こちらの小委員会ではこれまで、デジタル言語資源の在り方に関しまして、言語コーパスの整備状況やその活用事例、また、デジタル言語資源による社会課題解決の試みについて、大規模言語モデルの最新の状況なども把握しながら、皆様に情報共有を図ってまいりました。また、昨年度は、さらに、各地域の言語資源ということで、方言についても取り上げていただきました。

現在、AIをはじめとしまして、社会の状況、そして特にデジタル化の状況というのが急速に進んでいるというところでございますので、こういった中で、社会の変化や課題、ニーズといったものを適時適切に捉えた上で、改めてデジタル言語資源の整備状況や、社会課題解決のための取組事例について、引き続き把握を進めていただければと思っております。さらに、そうした現状を踏まえた上で、今後の整備・活用に伴う様々な課題、そして信頼できる言語データとしての意義というものを改めて確認するということを通じまして、これからの国語施策に関する御検討をお願いできればと思っております。

具体的には、配布資料3の当面の日程でございますように、前期と同様、各分野で言語資源の整備・活用に取り組んでいただいている有識者の方々からの事例紹介というものを中心にいたしまして、定期的を開催をし、年度末にはその結果を取りまとめるという方向で進めさせていただければと思っております。

説明は以上です。

○相澤主査

では、ただ今の御説明に対して、何か御質問や御意見などがありましたら、よろしく願いいたします。よろしいでしょうか。

(→ 挙手なし。)

御説明ありがとうございました。

続きまして、前期の審議報告に至る議論の中でもEUのテキスト使用に関するガイドラインの制定などについて触れられておりますが、この中で、ドイツ語コーパス拡充の背景にどのようなことがあったのか知りたいという前川委員からの御要望を受けまして、配布資料4「ドイツ語コーパス(DeReCo)拡充の背景」がございますので、こちらについての説明を事務局よりよろしくお願いいたします。

○鈴木国語調査官

配布資料4を御準備ください。ドイツ語コーパス(DeReKo)の拡充の背景について御説明申し上げます。

この話は、前川委員から、ドイツ語のコーパスの拡充ということに関して、何かその背景になることがあったのではないかと御発言があったことを受けてのものです。その時には、EUのテキスト使用ですとか著作権の関係のことが何か影響しているのではないだ

ろうかというお話でございました。

以前、ドイツ語のコーパスについてお話を伺いました広島大学の今道晴彦先生に、最近のドイツ語のコーパスの拡充の事実ですとか、その背景になる著作権等の動きについて御存じのことをとお尋ねしたところ、詳しくお答えいただきましたところを基に、配布資料4は作成しております。

ドイツ語のコーパスの拡充につきましては、2025年に、ドイツ語研究所から出ているドイツ語のコーパスに関して大幅な拡充がございました。2024年1月の時点で576億語ぐらいの規模だったものが、2026年1月には638億語の規模のコーパスとなっております。この期間で62億語ほど増えているということで、この増え方というのは例年になく非常に大きな幅で増えたということ、まずこういう事実がございます。

では著作権関係で何か大きな変化はありましたかというお尋ねをしたんですけれども、法律に関してこの分野に関わるような大きな変化が、この期間、24年の頃から最近までであったのかというと、思い当たるところはないですねというお答でございました。

ドイツの場合、著作権に関しましては、テキストデータマイニングに関する権利制限というものの規定が、2013年に10年間の時限的な規定として定められたという事実がまずあります。ただ、10年間の時限的な規定だという形だったのですが、10年たつ前、2017年に、これが時限的な規定ではなくて時限が外れた形での規定となって、翌2018年に施行されております。タイミングとしては、2018年、この頃というのは日本の著作権法の権利制限の規定が改正されたタイミングとちょうど同じ時期でございます。そして、実はその翌年、2019年に、EUから、EU全体にわたっての著作権、著作隣接権に関する扱いについて統一的な形にしようという指令が出て、各国でそれを法律に落とし込むということが行われたようでございます。ドイツの場合はこれが2021年に法律という形になりまして、より明確にテキストデータマイニングというところで受ける権利制限がはっきりと示されるようになってきたということは、背景としてございます。

ですから、ドイツの場合、日本と著作権の部分の扱いについて大きく違うのかということ、基本的にはかなり似た形です。ただ、研究分野においての複製等に関しては、日本よりもかなりデータ提供というものがしっかりと求められるという形ですので、その点においてはドイツの方が、日本の法律と比較する限りにおいて言うと、使うということに関してはより強い形で表現されているように思われます。ただ、それ以上に何かしらの改正が行われたかということ、そういうことはございません。

では、なぜこれだけ多く増えたかと言いますと、拡充されたデータの内容がポイントとなっています。拡充データとして何が入ったのかということはレポートが上がっておりまして、それを見ますと、ウィキペディアのユーザーによるオンライン上のディスカッションのデータというのが含まれております。これが非常に大量のもので、これが入ったことによってデータが大幅に増えたということだと、それ以外には考えられないだろうというのが今道先生の御見解でございます。

こうした拡充等の流れの背景には、ドイツの政府の大規模プロジェクトというのがございまして、データ基盤を整備していくという、その中において、ドイツ語に関するデジタルデータを集約して研究利用できる基盤を構築するText+という取組が始まったことによって、比較的予算がしっかりと確保されているということが背景にはあるという御指摘も頂きました。

御参考として、2枚目のところには、毎年度の予算－公的な予算(国だとか州からの予算)よりも、はるかに外部資金の方が多いということが特徴的な点かと思われます。

つまり、国としての政策的な方向性があることによって、予算がある程度確保されていたということ、そして、ちょうどタイミング的に膨大なデータを持ったものを取り込むことができたということが、ドイツ語コーパスのここ最近の大幅な拡充につながっていたということになろうかと思えます。

説明は以上でございます。

○相澤主査

ありがとうございました。法律面、そして内容的に、ウィキペディアユーザーによるオンライン上でのディスカッションのデータによる拡充量の大幅な増加、それから国家研究データ基盤としての政府主導、あるいは外部資金のサポートを受けたプロジェクトという3点についての御説明でございました。御質問や御意見等ございましたら、よろしくお願いいたします。

○石川副主査

よろしいでしょうか。

○相澤主査

お願いいたします。

○石川副主査

大変貴重な資料を共有していただきありがとうございます。

こうした資料を拝見しておりますと、この小委員会において、今後、諸外国における言語資源構築への公的支援の実情を調査することも重要なポイントの一つになるのではないかと思います。本邦においてはBCCWJ2など、国の支援で大型コーパスの構築がなされているわけですが、諸外国はどのような助成を行っているのか、あるいは公的支援よりもむしろ民間等から資金を集めているということであるなら、それはどういう枠組みで動いているのかといったことについて、基礎的な情報を整理し、今後の日本の言語資源開発支援に役立てていくことができるとよいと感じた次第であります。

貴重な報告、感謝申し上げます。ありがとうございました。

○相澤主査

ほかはいかがで。前川委員、よろしくお願いいたします。

○前川委員

貴重な報告をありがとうございました。ちょっと私から補足させていただいてよろしいですか。

○相澤主査

お願いいたします。

○前川委員

これは画面の共有ってできますか。

○相澤主査

はい、大丈夫でございます。

○前川委員

では、ちょっと画面共有させていただきます。

グラフが見えていますでしょうか。

○相澤主査

はい。

○前川委員

これがIDS、今のお話で出てきたDeReKoというのを作っているインスティテュート・フュア・ドイチェ・シュプラッハというマンハイムにある研究所のホームページから取ってきた図なんですけれども、これがDeReKoの増え方なんです。これは私が2024年に先方にお邪魔したときに事前の勉強でいろいろ見たときの資料を今出しているんですけど、今日のお話はこの先の……。数は見えていますか。

○相澤主査

はい、数は見えています。

○前川委員

このグラフは2024年までで、この後の変化についてということで御説明いただいたと思うんですけど、実はここで物すごく増えているんですね。これについては先方で、どうしてこんなに増えたんですかということ、私、先方の担当のWitt先生という方がおられて、あと、向こうの所長のRobin先生という方からいろいろ教えていただいたんですけど、この時点で結局、EUのデータ提供に関する指示なんかが話題に上がっていて、その時点で非常に多くの出版社と有償のデータ提供の契約を結んだということでした。

非常に多くの出版社というのは、IDSのホームページをのぞきますと、全部出ています。アワ・テキスト・プロバイダーみたいなのがありまして、見ていただくと分かりますけれども、非常に多くのものが新聞社です。それと、一部の出版社。それと有償の契約を結びまして、年間一定の金額を研究所が払う。それに対して新聞とか、多分、雑誌とかも入っているんじゃないかと一余り詳しくは調べていないんですけども—そういったものをDeReKoに取り入れられるということで、ここで、これは多分、50億から250億ぐらいに増えていますよね。という形で、一気に200億ぐらい、ここからここで増えたというのが過

去に一度あって、今回、ウィキペディアの話を私は実は知らなかったんですけども、そういう段階的にいろいろとステップを踏んで増やしているんだということが、今日のお話でよく分かりました。

私の補足説明は以上です。

○相澤主査

どうもありがとうございました。

○石川副主査

貴重な資料をありがとうございます。御提示いただいた資料で、急増期の年号が少し読み取りにくいのですが、これは何年のことでしょうか。

○前川委員

いや、私も今、読めないなと思っていたんですけども、多分2013年ぐらいだろうと思います。

○石川副主査

だとすると、テキストのデジタル使用権利に関する時限法の始まった年ということになりますよね、

○前川委員

ですから、その議論、法律の前にEUのディレクティブという、EUが各国政府に命令するという段階がありますよね。その後、各国政府が法律を実装するわけですが、多分、そのディレクティブを出したか、出す直前かぐらいのタイミングじゃないかと思います。

○石川副主査

なるほど。ありがとうございます。

○前川委員

では、共有を停止します。

○相澤主査

大変貴重な情報を頂き、ありがとうございます。2025年についても、2013年についても、大変有用な情報であると感じました。

○前川委員

今度、ちゃんと調べておきます。

○相澤主査

そのほか、よろしいでしょうか。

(→ 挙手なし。)

では、御説明、御議論、ありがとうございました。

続きまして、今期最初の言語資源小委員会ということで、配布資料1にある名簿順に、自己紹介も兼ねて一言ずつ頂くということになっております。たまたまですが、あいうえお順に行きますと、私が最初で石川副主査が2番目ということです。他の皆様、心の準備をその間にどうぞよろしくお願いいたします。

では、私がトップバッターとなりますけれども、相澤でございます。国立情報学研究所というところに長年おりまして、専門は自然言語処理や情報検索、今風に言えばAI、あるいは大規模言語モデルを専門としております。ふだん大量のテキストをデータとしてコンピューターで処理するという観点に慣れているため、恐らくこの小委員会を最も新鮮に思っているのは私ではないかとも感じるわけですが、一方で、文の構造ですとか、あるいは人と人のコミュニケーションのツールとしての言語も、言語を計算機で扱う上で重要なテーマで、皆様の様々なお話を伺いまして、毎回、大変勉強させていただいている次第です。微力ながら本小委員会に少しでも貢献できればと思っておりますので、どうぞよろしくお願いいたします。

では、石川副主査、よろしくお願いいたします。

○石川副主査

石川でございます。神戸大学に勤めておりまして、専門はコーパス言語学です。

御承知のように、言語資源小委員会は、文化審議会国語分科会の中に設置されているわけですが、実は、このこと自体が非常に重要であり、また、関係者として有り難いことだと思っております。今後は、この小委員会の枠組みをうまく使って、言語資源が幅広い国民にとって身近なものとなり、その有益性が広く認知され、実際に国民生活の向上につながっていくよう何らかの提案なり、提言をここから出すことができればと考えている次第です。どうぞよろしくお願いいたします。

○相澤主査

では、続きまして、植木委員、よろしくお願いいたします。

○植木委員

ありがとうございます。同志社大学の文学部に所属しております植木と申します。

私自身は古典文学、日本の中世文学に興味を持って、それを教えている立場なので、現代語のコーパスについてはお役に立てるところが少ないと思うんですけれども、ただ、今の言葉もいずれ古語になっていくということで、言葉を記録していく、蓄積していくということが非常に重要だということは、古典を勉強している立場からも強く思うところではあります。私自身は大変微力ですけれども、この場で本当にいい学びの機会を頂いているので、学ばせていただいて、議論に参加していければと思っております。どうぞよろしくお願いいたします。

○相澤主査

ありがとうございます。

では、神永委員、よろしくお願いいたします。

○神永委員

辞書編集をしております神永と申します。よろしくお願いいたします。この言語資源小委員会は、私は3年目になりまして、辞書編集を長年やってきたものですから、辞書編集の立場からコーパスをどのように活用するかといったことを考えたりとか、発言させていただいたりとかいうことをしております。

私は元々『日本国語大辞典』の第2版を担当しておりました。この『日本国語大辞典』は、

実は2032年に第3版を出すように、今、編集の活動を行っております。もう私は会社を定年になりまして10年たつんですけれども、いまだにちょっと編集のお手伝いをさせていただいているということです。実は、この『日本国語大辞典』というのは、初版から用例が命だと考えております。第2版までは特にコーパスということを活用したというわけではなかったんですけれども、今度、2032年に刊行されます第3版ではコーパスを大いに活用できるようなことになってきておりますので、実際にそれをどのようにして使っているのかということも併せて御報告させていただければ—今までも少しずつやってきているんですけれども—今期におきましてそれがもうちょっと具体的な形になっているはずですので、それができればよろしいかなと思っております。

コーパスというのは用例や語の用法の目安でありますし、そういった意味での宝庫でもありますので、そういった意味で大変勉強させていただいておりますし、今後も楽しみにしたいと思っております。

それから、昨年度から方言についていろいろと議論がなされておりました、私も実は、私ごとなんですけれども、『日本国語大辞典』で方言欄を担当しておりました、膨大な方言資料、方言データというのが存在している、でもそれがもうばらばらになっているということもよく知っているんですけれども、それがこの言語資源小委員会の中でどのような形でまとめられていくのかということも実は大変興味がありまして、すぐには無理かもしれませんが、できれば何か方向性が見いだせればよろしいかなと考えております。また今年度、よろしくお願いいたします。

○相澤主査

よろしくお願いいたします。

では、日高委員、よろしくお願いいたします。

○日高委員

関西大学の日高と申します。今期初めて委員を拝命しまして、小委員会での過去の議論の資料を追いつけながら、今期の課題についてもいろいろと想像しながらお聞きしているところです。専門は、方言を含む現代日本語の変種に関する研究に取り組んでおりまして、今は関西大学に所属していますが、その前は14年ばかり秋田大学におりまして、秋田の方言収録事業などにも携わってきました。そういった地域言語の収録・保存などに関

心を持って取り組んでいます。

それから、神永委員とも関わりがありますけれども、現在進行している『日本国語大辞典』の第3版に向けての改訂の編集委員も拝命しておりまして、方言の項目の改訂に取り組んでいるところです。恐らくこの小委員会でも言語資源として方言を含む日本語の多様な変種を扱うということを想定した議論が進んでいくようですので、それに関して幾らかでも私の経験が反映できるように、議論に参加できればと思っています。どうかよろしくお願いします。

○相澤主査

よろしくお願いいたします。

では、前川委員、よろしくお願いいたします。

○前川委員

前川でございます。国立国語研究所で所長を務めております。ということで、文化庁からは、委託事業を受けている立場なので、どうもいつも何か言うときに利益相反が気に掛かって、歯切れの悪い言い方しかできないんですけれども、ずっと言語資源の開発に国語研で携わってきた人間です。いろいろと苦勞もしてきましたし、それから、先ほどのようないろいろな知識というものも身に付いている部分もありますので、微力ながら何かお力になれたらと思っています。

一つだけ、今後の運営について希望を述べさせていただくと、これまでの審議は主に大體書き言葉関係の言語資源ということが念頭にあったと思います。それはこの小委員会だけではなくて、世の中全体がそうなんです。大規模言語モデルというのはほとんどは書き言葉のデータなわけですが、人間にとって本質的な言語というのは多分、話し言葉の方であって、書き言葉というのはあくまで二次的に派生してきたものだろうと思います。そういう意味で、話し言葉のデータみたいなものの言語資源というのも非常に重要で、言語研究はそうでもないのかもしれませんが、その隣接領域にとっては実は話し言葉のデータというのは非常に重要なものですので、そういったものにも視野を広げていければと思っています。

以上です。

○相澤主査

ありがとうございます。

では、森委員、よろしくお願いいたします。

(回線の状況が不調で、音声が出ず。)

では、順番を交代することにいたしますね。

森川委員、よろしくお願いいたします。

○森川委員

Gakken辞典編集部の方と申します。よろしくお願いいたします。今期から初めて委員を拝命いたしましたので、やっていきたいと思っております。

私は、名簿の中にあるように、日本書籍出版協会の文化庁の行政に対する委員会の副委員長を務めております。また、同時に日本雑誌協会の、同じく文化庁の行政に対する委員会の委員長も務めておりまして、両協会の推薦を受ける形でここにおります。ですので、両協会の動向などを踏まえた意見を申し上げることもあると思っております。また、出版業、あるいはその中の編集業、特に辞典編集に関して、神永委員と同じように長年続けてきましたので、そういうところからの御意見を申し述べたいと思っております。

辞典を作るに当たって、コーパスというのは非常に有用なものですし、コーパスが充実することで、より良い辞典というのが作っていただけると思います。それと同時に、そういう作り上げた辞書というものが版元横断で一発検索できるような—アメリカなんかではそういうサイトもありますけれども—そういう日本語辞典のサイトみたいなものも作れると、両輪のような形で発展していただけるのではないかなというところもちっと誇大妄想みたいな形で考えておりますので、そういうことも踏まえて、皆様の御意見等をいろいろ聞かせていただくのを楽しみにしております。どうぞよろしくお願いいたします。

○相澤主査

ありがとうございます。

では、森委員、よろしくお願いいたします。

○森委員

大変失礼いたしました。日本テレビの方でございます。委員に加えていただいて2年目

になります。今日は文化庁に伺えなかったんですけども、いつも文化庁に伺って、前後の時間で先生方に質問をするのがとても楽しくて、また有り難い機会を頂いているなど思っております。

アナウンサーとしてはもう30年になりまして、ここ数年は、新聞と放送の各社が集まる用語の委員会にも名前を加えていただきまして、そちらの会議にも出ております。最近で言いますと、例えば小学生の子供が事件・事故に関わった際に、敬称が「さん」なのか、「君」なのかという問題であるとか、あとは「魚雷バット」というのをきっかけとして、野球の言葉には何かと物騒な表現が多く出てきますね。例えばアウトが「死」－死ぬという漢字を使われていたりとか、その辺りを、いい悪いではなくて、どうでしょうかという議論をする場を持っております。こちらでお話を聞かせていただいたこともそちらの委員にも共有しておりますので、今後ともよろしくお願いいたします。

○相澤主査

ありがとうございます。どうぞよろしくお願いいたします。

では、このタイミングで、もし追加で何かコメントや、これだけは言っておきたい等ございましたら、お願いしたいと思いますが、いかがでしょうか。よろしいですか。

(→ 挙手なし。)

では、特にないようでございますので、最後に事務局から連絡をよろしくお願いします。

○鈴木国語調査官

承知しました。

配布資料3に、当面の予定という形で日程をお示ししております。次回の言語資源小委員会は、第14回ということで、6月30日火曜日でございます。今日と同じく15時から17時の予定で、有識者による事例等紹介と書いてございますが、予定しておりますのは、国立国語研究所のBCCWJ2の進捗のお話と同時に、先ほど前川委員から御発言のあった話し言葉のコーパスについても併せてお話いただくことを予定しております。それ以降の回についても、既にどなたにお話いただくかというところは大体調整が付いているところでございます。また楽しみにしていただければと思っております。

以上でございます。

○相澤主査

ありがとうございます。

では、これをもちまして本日の言語資源小委員会は閉会といたします。ありがとうございました。