



HOKKAIDO UNIVERSITY

Title	言語データの継承と保存に関する課題について
Author(s)	加藤, 重広
Citation	北海道大学文学研究院紀要, 161, 35(左)-49(左)
Issue Date	2020-12-18
DOI	https://doi.org/10.14943/bfhhs.161.l35
Doc URL	https://hdl.handle.net/2115/80045
Type	departmental bulletin paper
File Information	1004_bfhhs.161.l35.pdf



言語データの継承と保存に関する課題について

加 藤 重 広

本論は、広義の言語研究の一環として収集されたデータや関連資料がどのような危機にあるか、また、それらのデータや資料（以下、言語データ）を継承・保存する上でどんな課題があるか、また利活用に際して遵守すべき研究上の倫理とはどのようなものなのか、について論じるものである。

1. 現状認識と課題の設定

言語学における「データ」の位置づけはそれぞれの枠組みによって大きく異なるが、「データ」に基づいて仮説を検証する手続きを踏んで結論を得ることから、自然科学だとする見方もある¹。一方で、記述言語学では特定の言語変種（linguistic variety）を捉えるのに、その言語や方言を母語とする情報提供者²から言語データの提供を受け、記述する。このとき、研究者の分析が

¹ 自然科学とする考え方は理論言語学で根強い（中井, 1997, 福井, 2001, 2012）。理論的研究の中では、例文を比較することによる「実験」もおこなわれるが、実験の中には文の文法性の判断が研究者自身の内省によるものもあるため、「実験」とも「自然科学の手順」とも言えないとする批判も多く見られる。現に、結論を補強するために、内省判断を恣意的に決めた研究も以前は実際に散見された。近年では、その種の問題点を踏まえて、内省が論証の唯一の根拠とならないように慎重に手順を決めたり、複数の被験者の内省を利用する社会言語学的方法論を取り入れた研究手法をとったりするなど、状況が変化してきた（加藤, 2019）。

² かつては「インフォーマント」と呼んでいたが、informant に「密告者」といった意味があることから、linguistic informant とする呼称も見られるものの、近年は linguistic helper や consultant あるいは collaborator などと呼び方が変わってきている（第3節も参照）。

入っているデータと分析以前の生データがあり得る。前者の例としては、例えば、IPA での音声表記をしている場合や音素表記をしている場合がある。IPA での記述はグローバルな共通性はあるものの、分析者の音声学の知見によって精度や確度が異なる上、どの時代の IPA で記述しているかによっても精度や解釈が異なる。音素表記してある場合も、どういう音韻体系を指定しているかが分からなければ、データを正確に理解することは難しい。分析の入っていない生データとしては、母語話者が発音した単語などの音声などがあるが、それには、もちろん、口承文芸や個人の思い出語り、複数の母語話者による会話などの音声データ（ときには、動画データ）などがある。しかも、これらのデータの形状や形式は千差万別である。

もちろん、論文や報告書、あるいは言語資料として公刊されているものは研究成果であって、言語データとは異なる。研究成果とそのもとになった言語データの同一性が高い場合も当然あるが、研究者は収集したデータをそのまま公刊することは少ない。調査結果を言語資料として公刊する場合も、文字化するために音声表記や音韻表記をしていることが普通で、付録として音声データを準備していることもあるが、以前は希有なことだった。

問題は、これらの多様な言語データ（公刊される前のもの）が、その収集を担った研究者の退職や物故によって、失われてしまう可能性が高い現状である。言語音の録音は、蝸管や円盤の蓄音機などが発明された 19 世紀後半から可能になったため一世紀半に及ぼんとする歴史があるが、言語研究者が個人として録音機材を使って言語音を録音して残すようになったのはおおむね 1960 年代であり、現時点で半世紀を超えている。テープレコーダーの原型も 1930 年代に開発されているが、業務用でなく、個人が使用可能なものが出回ったのは 20 世紀の後半である。しかし、記録用のテープが高価で貴重だった 1950 年代の音声データはそれほど多く残っているとは言えない。1962 年にカセットテープ（いわゆるコンパクトカセット）がオランダのフィリップス社によって開発されると、オープンリールに徐々に変わっていったが、カセットテープが廉価になり入手しやすくなった 1970 年代には現地調査する研究者の多くが音声を録音したデータを残すようになった。この時期

はオープンリールとカセットテープが混在しているが、全体的に見て、ほぼ半世紀前から音声データが増えたと言えるのである。

半世紀前に録音した言語データを残し始めた研究者の年代や年齢はさまざまであるが、試みにこの時期に研究者としてのキャリアを始めた研究者像を想定してみよう。このような研究者であれば、二十歳代から三十歳代にかけての時期に調査データを取り始めたと思われるが、個人差が大きいので、30歳で音声データを取り始めたとして想定してみる。半世紀前に30歳であれば、現在は80歳である。詳細は、このあと論じるが、現在の日本人の平均余命を考えると、仮に想定したこの人物はおそらく10年程度前に定年退職しており、言語データを個人的に所蔵している可能性が高い。統計的な平均値だけから推定すると数年のうちに鬼籍に入るとも考えられる。もちろん、平均値に基づく推定なので、もっと長命である可能性もあるが、それでも、現時点でこれらのデータの保存を始めなければ、永久に失われてしまうことになると考えるべきだろう。

現状認識の詳細は、このあと述べるが、少なくとも現時点で何か行動を起こさなければこれらのデータが永久に失われる危機的事態にあると筆者は考えている。この認識は、筆者の周辺の言語研究者にも共有されている。もちろん、危機言語でなければ、母語話者はたくさんいるので、苦勞して半世紀前のデータを保存しなくてもよい、と考えることもできる。しかし、記述言語学の重要性が理解されるようになったこの半世紀にあっては、相当数の貴重なデータがあると考えられる。加えて、現在は失われた半世紀前の言語態の記録があったとしたら、たとえ大言語のデータだとしても十分に価値はあるだろう。

いま失われつつある言語データを保存すべきだという現状認識が妥当なものだとすれば、私たちの課題は、どのように保存するためのしくみをつくるのか、ということであり、それが今後の言語研究に利用できるようにするためにはどのような問題があるのか、ということになる。そして、言語データは、特定の言語話者により提供された個別データの集積である以上、それを利活用する上で、どのような研究倫理上の問題があり、どのようなことに

意を用いていくべきなのか，ということも課題になる。

2. データの形状と形式

データと一口に言っても，その形態・形状は多様である。例えば，文献資料などは研究者が作成したものではなく，第三者が過去において作成したものである。文献を使って行う研究における「文献」はそれ自体が原資料の場合もあるが，貴重な資料であれば複製情報が用いられることもある。また，碑文のように原資料を移動しにくい場合も複製情報を用いることになる。

また研究者が収集作成した資料（＝第一次資料）もあり，現在の言語研究が直面するデータ喪失の問題で採り上げているものの大半はこれである。前節では，以下のような大まかな区分を立てて説明をした。

表1 データ形式の区分

研究資料	研究成果を作成するための資料	第一次資料	研究者の分析・判断を含まない資料
		第二次資料	研究者の分析・判断を含む資料
研究成果	論文や報告書など成果として公刊されたもの		

例えば，母語話者の音声を録音したカセットテープは，「研究資料」であるが，研究者の分析や判断を含まないので「第一次資料」だと前節では述べた。これに対して「音素を決めるためのメモや音素一覧を含むフィールドノート」は「研究資料」であり，「第二次資料」だということになろう。

この「一次性」に関する区分は大まかな区分としては有効であるが，研究者の分析や判断がどの程度関与するかは容易に決められないものである。例えば，ある単語リストを母語話者に読み上げてもらい，その音声を録音する場合，リストの単語の配列に研究者の判断が全く反映しない，無作為のものである可能性は低い。特定の音声あるいは音素を含む語をまとめたり，音節数を基準にしたり，音声構造や，アクセントや音調の共通点で括ったりする

ことは十分にあり得ることであり、その場合は、分析で得た知見をもとに判断していることになる。

本論で想定している言語データは録音資料と呼ぶべきものであるが、実際は音声録音されているテープだけであっても、活用の余地はあまりない。多くの場合、録音内容がわかるようになっているフィールドノートやメモの類いがあるはずである。少なくとも、単語の音声リストであれば、簡易表記や意味の一覧があるはずであり、例文の読み上げ音声も対応する文の一覧があるはずである。いわゆるテキスト（民話や伝承の独話や対話など）はすぐに文字に起こせない可能性があるので、音声だけのケースが考えられるが、梗概や題目は残してあることが多い。

音声資料の録音スタイルも研究者によりさまざまである。言い間違いや読み間違いがあっても止めずにそのまま録音し続けたり、途中で別の話が始まっても中断することなく続けたりする場合もあれば、誤りや言い直し、ノイズを避けるため、録音し直すこともある。単語リストと対応する音声だけがあれば混乱はないが、言い直しがある場合、いずれが言い間違いなのかあとから判然としなくなる可能性がある。

他の形態のデータとしては、ノートやメモ、カードなどの書記資料もある。フィールドノートは、分析上気づいた事や野外調査の予定やメモを含む場合もあれば、必要最小限の言語学的な情報しか含まないこともある。ノート類は、現在スキャナーを利用してPDFにして保存することが容易になっている。付箋やメモ、挟み込まれた別葉の資料などはうまく残せないこともあるが、おおむね、丁寧に作業をすれば逸失は防げるだろう。問題は、これらの書記資料と音声資料の対応関係がどの程度明らかなのかということにつく。

書記資料とは異なるが、写真や図、地図などの図画類も研究資料として残されていることがある。地図は、現地調査にかかわる記録や簡便な言語地図である場合もある。これらは、直接的な「言語データ」ではないが、言語データを補完するものであり、無視することはできない。加えて、書記資料以外の「関連資料」も見られる。その多くは、民族誌にかかわる物質的資料で、

儀礼の用具や民族衣装に代表されるモノの資料である。

古い時代の写本などはもちろんいずれも「一次資料」だが、ほぼ同じ内容の複数の写本があればいずれが原典でいずれが複製かという関係を明らかにすべきことも多い。しかし、研究者から見れば、これらはいずれも「一次資料」である。一次資料をデジタル化したものや写真撮影して一冊の本にしたものなどは一次資料の複製物に相当する。写本を翻刻したものは研究者の解釈を含むので二次資料となる。

言語データはここまで見ただけでも「関連資料」を含め、非常に多様であることがわかる。書記資料は、ノートやカード、あるいはメモといった違いはあるが、判読できれば資料としての価値は失われない。しかし、音声資料は前節で言及したように、カセットテープが多いにしても、オープンリールもあり、近年ではCDやDAT、MDなどがあった。しかも、これらにはいくつかの規格があり、すべてが同一の環境で利用できるとは限らない。例えば、標準的なカセットテープはノーマルポジションやハイポジションなら再生できるが、メタルテープは対応していない場合があるとも言う（多くの場合は、筐体の形状の問題なので再生そのものは可能である）。機材さえあれば、ひどく劣化していない限り、再生は可能であるが、今後長期的に利活用していく場合を考えると、デジタル化して簡便にデータを取り扱えるようにして置くことが望ましい。しかし、デジタル化する場合にも、wav なり、mp3 や mp4 なり、どのようなデータ形式にしておくのが望ましいのか検討しておく必要がある。現状のデータ形式が今後も長く使われ続けるとは限らないことを考えると、継承し、保存するには、データ形式を変換する必要性が生じると思われる。音質やその他の特性で優れていても、特殊な形式でデータ変換が容易でなければ、保持することが難しくなったり、手間がかかったりすることになる。

Himmelman (2012) では、原データ (raw data) と主要データ (primary data) と構造データ (structural data) にわけ、構造データを二次データ (secondary data) とも呼んでいる。上記で筆者が一次資料と呼んだものが「原データ」におおむね相当し、二次資料と呼んだものが「主要データ」に相

当するようだ。

表2 Himmelman (2012) によるデータタイプ区分

	歴史言語学データ	現代のデータ	メタ言語的技能に基づくデータ
原データ	碑文	音声記録・画像記録	個別の反応や評価の記録
主要データ	翻刻	フィールドノート・対訳付きテキスト	実験の統計分析結果・フィールドノート
構造データ	音法則	グロスつき例文の記述・辞書・データベース	処理速度・頻度データ

更に Himmelman (2012) は歴史言語学データ・現代のデータ・メタ言語的技能に基づくデータにわけ、3つのレベルの例を挙げている。その概略を表にしておく³。紙幅の都合ですべての情報を盛り込んではいないが、概ね本論の区分と一致していて大きな相違はないように思う。あるとすれば、本論で言う「一次データ」が「原データ」で、「二次データ」が「主要データ (primary data) = 一次データ」にあたり、用語法のずれがあるくらいだろう。ただ、本論では「一次データ = Himmelman (2012) の raw data」と「本論の二次データ = Himmelman (2012) の primary data」に質的差異はあるものの、峻然と境界があるのではなく、両者が混じっていて連続的に分布する関係にあると考えているので、別の機会に詳しく論じることにした。

3. データにかかわる権利

研究者が収集したり作成したりするデータは誰のものだろうか。多くの場合、研究者個人が保有している以上、研究者個人の所有物であることに異論

³ Himmelman (2012) では、3つの表(表2 : p.192, 表3 : p.193, 表4 : p.196)で示されているが、引用者が取捨選択してまとめ直した。

の余地はない。しかし、これはカセットテープやフィールドノートといった所有物に関する所有権のことである。

例えば、学会が刊行する雑誌・学会誌に掲載された論文に関する権利は、特定の誰かが占有しているわけではない。一般的に、論文の内容に関する権利は著者が所有しているが、学会誌という印刷物になった場合版面をつくるための編集作業が必要なので、版面の権利は出版社か学会が持っていることが多い。また、学会には公衆送信権があり、多人数の会員に公刊された論文を届けることができる。また、近年は論文のフリーアクセスも増えているが、学術成果を広く社会に発信して、有効に活用してもらうためにおこなっているのだろう（日本における言語研究にかかわる学会の多くは任意団体であるが、社会貢献の責務を設立の趣旨に謳っていることが多い）。このような学会誌の掲載論文は権利関係が比較的単純でわかりやすい例である。しかし、研究者個人の作成・収集したデータとなると複雑で、難しい。

研究者個人が、分析や論考の途次に得た着想やアイデアを記したメモであれば、その内容もメモそのものも研究者個人のものであろう。では、ある言語を調査するために現地を訪れた研究者が母語話者に依頼して発音してもらった音声データはどうだろうか。協力してくれた母語話者は研究者個人の研究のために発音や吹き込みをしているので、研究者個人がその研究上で分析に用いる場合は、研究上の権利侵害は発生しないと考えられる。しかし、研究者がその音声データをなんらかの方法で広く公開するとしたらどうだろうか。研究者とインフォーマントが詳細な契約書を取り交わしているケースは少ないと思われる。例えば、母語話者が提供した音声を学会といった聴衆が特定される（所属会員しか参加しない）場で事例として聞かせるケースはどうだろうか。実際にそういう事例はあり、調査時に学会発表で聞かせることを考えていれば、言語研究者はインフォーマントに口頭でそのように活用する旨伝えて許可を得ていることはあるだろう。しかし、それを書面で「許諾範囲の特定」として定めているケースは寡聞にして知らない。

学会発表で研究者は直接的な経済的利益を得ることはまずない。しかし、学術講演のなかには（日本では少額であるものの）謝礼をもらうケースも多

い。おそらくその場合もインフォーマントにその謝礼の一部（数円から数十円程度にしかないだろう）を渡すことは考えがたい。多くの場合、言語研究者とインフォーマントの間には信頼と友誼に基づく人間関係があり、金員の授受は回避される傾向にあるからだ。実際に、亀井・河野・千野（1996）の「現地調査」の項目でも、日本ではインフォーマントに調査協力の対価として現金を支払うことは少なく、長期的に調査協力を得ようとする場合には避けた方がよいという趣旨の記述がある（その一方で海外では現金を支払うケースもあるとしており、実際に、現在の言語調査では現地の協力者に主に時給で計算して現金を対価として支払うことも珍しくはない）。しかし、日本国内での方言調査などでは、手土産⁴などを持参することはあっても、現金を渡すケースは現在でも多くないようである。これは、インフォーマントは学術調査への協力といった社会貢献をしているという見方があり、対価の支払いは社会貢献ではなく一種の労働という扱いに評価を下げてしまうことになるからだろう。

Eckert（2014）は、協力者に対価を払うことについて触れている。その都度、支払う方法もあれば、一定期間ごとに支払う方法もあり、例えば、毎年現地調査に行く場合に、特定のインフォーマントに協力を得ることを確約しているのであれば、長期的に対価を一定額ずつ払う方法もある。ただし、協力関係が長く続けば言語知識の提供とその対価だけにとどまらず、他の知識の提供や移動の便宜から生活物資の提供にまで及び、ビジネスか友人関係か境目が決められないような対人距離になってしまうと Eckert（2014：23）では記す。

研究者と調査協力者の個人的な信頼関係が成立しているうちは、多くの場合、問題になりにくい。問題が生じるのは、その一方または両方が物故して

⁴ 現地調査の方法論に触れた文献の中には、謝礼代わりに手土産を事前に用意すべきことや手渡すタイミングなどに言及しているものがある。乾物などを除き食品は避け、長期保存が可能な日用品として石鹸や洗剤、あるいは、調査者の大学名の入った文房具類や記念品（キーホルダーなど）が適しているといった知識は、以前であれば、現地調査に同行した初心者先輩が伝授するものでもあった。

しまった場合だろう。物故せずともその権利と義務を当人が管理できない状態になれば同様の事態と見ていい。

本論で想定しているデータのうち、研究者自身が作成したもの（語彙リスト、例文、文字起こししたテキスト、分析メモ、着想のメモ、調査の記録など）は、研究者自身の権利物だと、考えられるので、研究者自身が提供や譲渡を了解すれば特段の問題はない。もちろん、その場合には、研究者と提供・譲渡を受ける個人や組織の間で覚え書きか契約書のような書面を取り交わして残しておくべきだろう。研究者が物故したり権利管理ができなくなったりした場合は、遺族ないし代理権者と書面をかわすことができる。そのときに重要なことは、研究上のデータを管理する方法と公開や提供の範囲をあらかじめ決めておくことである。

例えば、ある言語・方言（ここでは仮に X 語とする）の言語データに興味を持つのは概ね専門家だと考えられるので、研究に活用してくれる研究者に提供できるチャンネルがあればよく、無制限に公開する必要はない。海外には、誰でも言語データ（を含む民族関係の調査データ）を提供でき、誰でも（登録などは必要だが）利用できるデータベースもあるが、典拠や収集方法が不明確な言語データが制約なく提供される状況は、そのデータの質の保証もなく、無用な混乱を惹起しかねない。

権利上の問題を考えなければならないのは、インフォーマントに提供を受けた音声データなどである。もしも、インフォーマントが X 語の単語を発音したのだとすれば、その単語は著作物ではないが、一種の実演ではあるので著作権法第 89 条に言う「著作隣接権」での保護対象にはなり得る。研究者が提供を受けた音声データを配列したり、取捨選択したりして語彙集・語彙リストを作成すれば、著作権法に言うレコード製作に相当する著作隣接権が認められるだろう。

もちろん研究者が音声データをまとめても、インフォーマントが音声データを提供しても、それらは学術目的であり、営利目的ではないので、権利料などは想定していない。インフォーマントの実演による著作隣接権は、保護の対象なので、不特定多数が利用できるようなデータベースに無断で登録す

るのは好ましくなく、著作権者による許可がなければならない。通常、研究者に提供した音声は研究成果になるプロセスで活用され、学術研究に資する目的で使われると想定されているが、それだけで自由な公衆送信権を許諾したとは考えられないからである。インターネット上などで公開する場合は、学術研究に資する目的で利活用されるとは限らず、悪意のある第三者が著作隣接権者たるインフォーマントの著作人格権を侵害することがないとは言えない。また、実際にはそれほど多くないかもしれないが音声だけでなく、ビデオ撮影などで動画データになっている場合には、肖像権にも配慮する必要がある。従来の記述言語学での調査では必ずしも動画は必要ないが、会話分析やマルチモーダル語用論などでは身体の動きや位置関係、視線や指さし指示などを含めて分析するために動画に記録することが増えている。

節を終えるにあたり、簡単に議論のポイントをまとめておこう。

表3 データ種と権利の関係

データ種・活用形態	権利者	権利形態
フィールドノート (メモ類含む)	研究者	著作権
音声データ	インフォーマント	著作隣接権 (実演)
	研究者	著作隣接権 (編集・作成)
動画	インフォーマント	著作隣接権 (実演)・肖像権
	研究者	著作隣接権 (編集・作成)
一般公開	研究者	公衆送信権

著作権と著作隣接権には人格権が認められる。通例、これらが学術研究にのみ用いられるのであれば、著作人格権や著作隣接人格権は侵害されないであろう。研究成果としての論文や発表などは批判の対象になることがあるが、根拠と論理を明示しながら科学的な議論として批判する限りは著作人格権の侵害にはならない。むしろ、著作人格権があるから研究成果が無制限に保護されることになれば、学術の発展は見込まれない。誤謬や誤解があつて

も結果として学術の発展に寄与することはあり得ると考えるべきである。また、研究の不備や問題点を著者の人格を傷つけることに直結させなければ批判そのものも問題にはならない⁵。

また、本論では深く論じないが、研究を成立させる資源という観点からの論考にも意味はある。ある言語 X に関する研究成果が日の目を見るとき、言うまでもなく研究者個人の資質と努力（エフォート⁶）が最も重要ではあるが、研究協力者としてのインフォーマントの資質（＝言語知識）と努力（エフォート）も不可欠であり、加えて研究費も必要である。調査や研究に必要な研究費は、個人が支弁することももちろんあるが、科研費や民間助成金、所属機関の研究費を充てるのが一般的だろう。

4. 言語データが失われないための方策

研究者のキャリアはさまざまであるが、おおむね大学院を修了して研究職に就くのが三十歳前後、六十歳代半ばから七十歳前後まで研究を続けられる環境にあると考えてみる。定年退職などで研究職を離れても、研究成果にまとめられる言語データは手元にあり、研究する能力もすぐに衰えるわけではないから十年程度は自分のペースで研究を続けられると思われる。八十歳代でも実際に元気で研究上の生産性を失っていない人もあるが、学会に出ることなくなって実質的に研究を続けているとは言えない状況にある人も多い。筆者の知る限りでは百歳でも学会に出席しているという方もあるものの、これは希有な例で、学会での役職の経験など学会との関わりによって個々に距離は異なる。それでも、定年後十年程度のうちに研究の割合が低下し、八十歳

⁵ 他の研究成果を取り上げた論文や書評においては度を過ぎた批判が見られる。表現のしかたによっては著作人格権の侵害とも言える場合があるのは事実だが、本論での議論の対象ではないので、指摘にとどめておく。誤りや不備があっても、研究成果そのものは尊重され、保護される権利があることを踏まえた批判でなければならない。

⁶ 精神論的な「努力」ではなく、能力を意識的に使って研究や作業を進捗させるための時間と考える。従って、関与する人的資源には一定の努力が認められる。

代では研究から離れる人の比率が増大するようだ。もちろん、それまでに亡くなれば、研究を続けることはできなくなる。

第1節で述べた懸念は、研究職を離れて一定の時日が経過した言語研究者 (retired linguist として RL と略記する) が自宅に保有しているであろう言語データの行方であり、RL が物故した後に保有していた言語データをどう処分するか、という問題とつながっている。現在、その種の言語データを引き受ける公的機関はなく、残された蔵書も大学図書館などは管理や所蔵スペースの負担から寄贈を受けないとも聞く。蔵書は古書店が引き取るかもしれないが、言語データは廃棄される可能性が低くない。ほこりをかぶった幾箱もの段ボール箱に入ったノート類やカード類、テープの類いに価値があると、専門外の人が考えないのは当然だろう。RL が亡くなったり、管理できなくなったりすれば、ゴミとして廃棄されてしまうのは自然の流れである。

現時点で八十歳代の RL は、だいたい 1930 年代生まれの世代だろう。少し広めに設定して、1940 年以前に出生した世代 (= 八十歳代) を想定すると、これは、研究者個人が録音機材を持って現地調査に(おおむね 1960 年以降に)行くことが可能になった世代(第1節参照)でもある。若い頃から言語音を録音しながら言語データを収集するスタイルで研究に従事していれば言語データの量はかなり多くなる。しかも、半世紀程度前の言語データとなれば、現在も調査可能な言語であっても、変化の比較ができるから研究上の価値はある。すでに話者が絶えたか、ごく少数しかいない言語であれば、なおさらである。

これらの言語データを公的に受け入れて保管するしくみが無い以上、廃棄されなくても、かつて指導を受けた後続世代の研究者が個人的に継承して保管する程度のことしかできない。しかし、何十年か後に継承した研究者個人も研究の第一線から引き、研究から離れるようになると、やはり同じ問題が発生する。とすれば、いずれかの機関か組織が責任を持って受け入れ、利活用できるようにするしくみが必要になるわけである。

受け入れる機関・組織に求められることは、まず永続性であろう。長期的に保管し、利活用するにあたって、保存保管する機関がなくなってしまうの

では意味がない。国策としてでも対応する必要がある。

言語データはそのまま倉庫に入れておくだけでは、どんなデータがあるのかが分からず利活用はできない（単に「失われたい」という意味はあるが、それでもデータの中身が分かっているなければ、管理していないことになり、適切に保管することは難しいだろう）。従って、言語データの中身について一覧表のようなものを作成する必要がある。RL 自身は自分のデータなので、他者にわかる形では整理していないことが多いが、一覧表のようなものを作成済みだとしても、外部の人間が理解できるものにして置き、実際のデータとの照合も済ませておく必要がある。データそのものは、デジタル化すればそのまま現物をとっておかなくてもよくなる。本来は、細部まで確認するために現物があつたほうがよいが、劣化も進み、保管の場所も必要になるので、デジタル化した時点で現物は RL か家族に返却することになるだろう。

5. まとめ

本論で述べた内容はいわば一般論であり、総論である。記述言語学的なデータは、いわば現地調査の数だけ存在し、それも研究者ごとに異なり、調査地や対象言語ごとに個別の事情があると考えられる。ある言語ではあまり問題にならない形態論の問題が別の言語では非常に複雑で記述上最も手間を要することもありうるし、音韻論が複雑でも統語論が単純という言語も相対評価の中ではあることだろう。

そうすると、われわれが実際にこの課題に対処するとして、一般論だけで制度設計をトップダウン的におこなっても、個別言語の記述の事情に合致しない可能性があることになる。当然、現地調査からの、個別言語からのボトムアップ的な設計思想も取り入れて、調整しながら安定的で「使える」しくみを構築しなければならない。

各論のパートも含めて論じ残したことは数多くあると思われるので、次の機会に詳しく論じたいと考えている。

参考文献

加藤重広（2019）『言語学講義』筑摩書房

亀井孝・河野一郎・千野栄一（編著）（1996）『言語学大辞典 第6巻 術語編』三省堂書店

中井悟（1999）『言語学は自然科学か：生成文法の方法論』昭和堂

福井直樹（2001）『自然科学としての言語学』大修館書店

福井直樹（2012）『新・自然科学としての言語学』筑摩書房

Eckert, Penelope (2014) Ethics in linguistic research, in *Research Methods in Linguistics*, Edited by Robert J. Podesva and Devyani Sharma, Cambridge: Cambridge University Press, 11-26

Himmelman, Nikolau P. (2012) Linguistic Data Types and the Interface between Language Documentation and Description, *Language Documentation & Conservation*, Vol.6, 187-207

* 本論文は、科学研究費補助金・基盤研究（B）（一般）（研究代表者：加藤重広，課題番号：18H00661）による研究の成果である。