

アイヌ民族博物館アイヌ語アーカイブについて

安田 益穂

アイヌ民族博物館ではこのほど、文化庁「アイヌ語のアーカイブ作成支援事業」の委託を受け、当館が所蔵するアイヌ語音声資料を中心にデジタルアーカイブを作成し、インターネット上で公開を開始した。URLは「ainugo.ainu-museum.or.jp」である。以下にその概要を紹介する。

1. アイヌ語アーカイブの成り立ち

(1) アイヌ語資料と語り手

アイヌ民族博物館では1976年の設立から2000年ごろにかけ、日高地方を中心に道内各地のアイヌ話者・アイヌ文化伝承者を対象に聞き取り調査を実施した。聞き手は博物館の学芸職員らで、調査内容は録音・録画テープに記録した。その後、伝承者は相次いで亡くなり、担当した学芸員もほとんどは退職、後には膨大な記録資料が未整理・未公開のまま残された。

(2) 資料のボリューム

音声資料は全体で670時間。内訳は、

- ①アイヌ語の口承文芸資料 103時間
- ②同じ物語を日本語で語り直したもの 34時間
- ③その他民俗調査 533時間

となっている。

またこれとは別に、儀式や芸能、建築、工芸技術などを記録した映像資料が約500時間分あり、これは現在整理中である。

今年度はそのうち、日高地方のアイヌ語の物語を中心に93件、約43時間分の資料を公開した。これは従来当館で出版した『民話ライブラリ』（A4判200ページ、CD2枚付）約20冊分に相当する。

(3) アナログ資料のデジタル化

デジタルアーカイブの作成には、アナログ資料のデジタル化が前提になる。ここでいうデジタル化は、ハードディスクなどに入れてパソコンで再生できるデータを指す。従ってDATやDVminiなど、再生機器が失われつつある前世代のデジタルテープもデジタル化が必要となる。基本的に音声はWAVとAAC(m4a)に、映像はMP4またはMOVで保存している。

先述の当館採録音声資料 670 時間はほぼデジタル化が済んでいるが、これはデジタル化が必要なメディアの約 20% に過ぎず、まだ 3000 時間分が残っている。民間博物館にこれをデジタル化する予算はなく、現在は文化庁の「アナログ音声資料のデジタル化事業」の支援に頼っているところである。

(4) 実施体制

当館採録の音声資料 670 時間分については、聞き起こしはほぼ済んでいる。これは 2011 年度から 3 年間、文化庁の補助を得て実施した。アイヌ語の口承文芸資料約 100 時間分の対訳はベテランが担当したが、アイヌ語混じりの日本語で語られている 540 時間分は若手 3 名が担当した。彼女らは当館で実施した若手アイヌを対象とする研修プログラム「伝承者育成事業」の修了生らである。本来は親から子、子から孫へと受け継がれるべきアイヌの伝承だが、録音を通じてとはいえ 3 年間毎日アイヌ古老の話を聞き起こしたことは、アイヌ民族の文化遺産をアイヌ自身の手に戻す、という意味でも意義があったと考えている。

彼女らの作成した文字データは、ベテラン学芸員や外部のアイヌ語研究者が校閲し、その成果の一部はその後「アイヌと自然デジタル図鑑」として当館ホームページで公開され、現在も人気コンテンツになっている。(ainu-museum.or.jp/siror/)

2. アイヌ民族博物館アイヌ語アーカイブの特長と構成

(1) 特長

次に、アイヌ民族博物館アイヌ語アーカイブの特長について述べる。

- ①「非破壊アーカイブ」……音声や映像などメディアファイルを細切れにせず、一本のファイルを丸ごとアーカイブ化した。
- ②全文データベース……聞き起こした全文文字データに対訳・注釈・文法情報などを加え、また音声・映像メディアとの連動を図るため行ごとにタイムコードをつけた。
- ③同期再生……カラオケの歌詞のように、現在再生している箇所を文字データ上でハイライト表示し、自動でスクロールする。
- ④ダイレクト再生……テキスト検索でヒットした単語を検索結果一覧画面でクリックすると、対応する音声や映像を直接再生する。また全文表示に遷移して任意の行から再生することもできる。
- ⑤多様な検索機能……単語検索、資料検索、辞書検索が利用できる。
- ⑥多様な表示方法……アイヌ語のカナ・ローマ字表記、逐語訳、グロスなどを表示・非表示することができる。
- ⑦二次コンテンツへ……その物語を原作とするデジタル絵本や挿絵など、二次コンテンツを表示する。

(2) 画面構成

基本的には単語や資料名などで検索し、結果一覧の中からどれかを選び、文字データを見ながらメディアを再生する、という流れである。

3. 利用の実際

(1) トップページ

任意のブラウザで URL に ainugo.ainu-museum.or.jp を入れると、トップ画面が表示される。(写真1)

▲写真1 トップ画面（検索画面を兼ねる）

① アイヌ語表記の選択

左横メニューの下にアイヌ語表記法の選択ボタンがある。アイヌ語にはカナ表記とローマ字表記があり、辞書や出版物も両派あるが、ここでどちらかを選べば既定値となり、その後はいちいち選ぶ必要がない。

② 検索方法

トップ画面は検索画面を兼ねている。

検索は大きく分けて、「単語を探す」と「資料を探す」の二つがある。単語検索では「辞典」「アイヌ語資料」「全文」のうちから対象を指定する。

写真2は「コタン」で単語検索した例である。

(2) 単語検索 (写真2)

The screenshot shows a search interface with two main panels. The left panel, titled '単語を検索' (Search for words), has a search box containing 'コタン' and a search button. Below it, there are two sections of search results: '検索結果 (全文から)' (Search results from full text) and '検索結果 (アイヌ語資料から)' (Search results from Ainu language materials). The right panel, titled 'アイヌ語全文表示' (Ainu language full text display), shows the selected result: '川上まつ子さんの民話(内) スズメの恩返し (1985)'. It includes a timestamp '12:21', navigation buttons, and a list of search results with their corresponding video thumbnails and titles.

▲写真2 単語検索と結果表示画面 (「コタン」で検索した例)

- ① 「検索結果 (全文から)」では、映像・音声資料の全文テキストから結果を表示する。「コタン」を含む行が106件ヒットし、スクロールすると全件確認できる。
- ② 同様に「検索結果 (アイヌ語資料から)」には667件、
- ③ 「検索結果 (辞書から)」には部分一致も含め23件ヒットしている。「コタン」は「村、集落」の意味だということがわかる。なお、辞書検索結果は一番下に表示されるが、ヘッダーにも同じ検索窓があって、こちらは別の画面を表示中でも辞書が引けるように、別ウィンドウに結果が表示される。

(3) 検索結果とアイヌ語全文表示

今見た三つの検索結果から、最初に②「検索結果 (アイヌ語資料から)」、すなわち口承文芸資料を見てみよう (写真2左下)。

行頭にスピーカーアイコンがついた行が並ぶ。スピーカーをクリックすると、その行の音声を画面の遷移なしに聞くことができる。「特長」の項 (2-(1)) で「④ダイレクト再生」と呼んだ機能である。スピーカーのアイコンではなく、行の任意の場所をクリックすると画面が左右2分割画面になり、右画面にその行を含むアイヌ語音声資料が全文表示される。

写真2の右画面には、「スズメの恩返し」という民話の選択箇所全文が表示されている。この画面にもスピーカーアイコンがあるのでクリックすると、現在再生中の行がピンク色にハイライトされ、5行ごとに自動でスクロールする。

ちなみに、「絵本」と書いてあるボタン (写真2右下) をクリックすると別画面が現れ、

この民話を原作とするデジタル絵本が YOUTUBE から表示される。

(4) 映像資料の全文表示

先ほど見た3つの検索結果から、今度は①「検索結果（全文から）」の行をクリックしてみよう（写真2左上）。

20000番台は映像資料である。フィルムアイコンをクリックすると、別画面で YOUTUBE から「コタン」の語を含む儀式の映像が表示された。映像資料も音声資料と同様、再生映像に合わせて文字がハイライト表示され、5行ごとに自動でスクロールする。

(5) 音声資料の全文表示

「検索結果（全文から）」を下にスクロールすると、30000番台は音声資料である。先述の通りスピーカーアイコンをクリックして音声と文字データを同期再生することができる。

(6) 資料検索

トップ画面に戻り、今度は資料検索を試してみる。

単語検索同様に①アイヌ語音声資料（口承文芸、会話、歌、祈り詞）、②音声資料③映像資料の3つから対象を指定することができる。

資料検索では、左画面に資料カード（目次）、右画面に全文を表示するような使い方が可能で、それぞれ当該箇所へジャンプして音声や映像を再生することができる。

なお、利用法の詳細はヘッダーに「ユーザーガイド」があるのでそちらも参照の上、実際に試していただきたい（写真1右上）。

4. 今後の可能性

(1) アイヌ語口演の復興へ

デジタル口承文芸アーカイブはさまざまな活用の可能性を持っている。その一つが実演からアーカイブ、アーカイブから実演への循環、すなわちアイヌ語口承文芸の復興への活用である。

アイヌ民族博物館では数年前から、隔週土曜の午後、伝統家屋の炉辺でアイヌ語の物語の実演をしている。「オルシペ アヌ ロー」（お話を聞きましょう）というこのイベントでは、聞き手の多くはアイヌ語がわからないので、最初に物語の内容を解説し、アイヌ語の語りの間、聞き手はレプニと呼ばれる棒で囲炉裏を叩いたり、配布されたプリントを見たりする。

ここにアイヌ語アーカイブを組み入れ、場面に合わせたイラストや日本語字幕をプロジェクターで映写し、その前でアイヌ口承文芸を実演することが考えられる。教室やホール、ステージなどでの活用に特に有効だろう。

(2) 新しいアイヌ語辞書へ

また、新しいタイプのアイヌ語辞書の登場が期待される。紹介した通り、単語を検索すると音声つきの例文がどんどん蓄積される。また現在もアイヌ語テキストは単語レベルでデータベース化し、言語学者が文法情報（日本語クロス）をつけ始めている。また、既刊のアイヌ語辞書を電子辞書化し、同じサイトで公開を開始した。今後言語学の優秀な研究者が加わることでこれらが統合され、今までにないアイヌ語辞書が生まれることが期待される。

(やすだ・ますほ／アイヌ民族博物館学芸員)

第41回大会シンポジウム「口承文芸デジタルアーカイブの課題と展望」

音声等を含む多様なデータベースの現状と課題 —人文情報学の立場から—

後 藤 真

本報告では、特に人文情報学の立場から、口承文芸に関わるデータベースの一般的課題について述べるとともに、国立歴史民俗博物館（以下歴博）が保有する民謡データベースの概要について説明するものである。

特に、本報告では、民謡や音声、口承文芸に関わるデータをいかに長期的にアクセス可能な状態とし、かつ研究するものとして使うことができるかについて述べる。この課題を整理するためのひとつの材料として、また情報の提供として歴博の「民謡データベース」を紹介したい。その後、人文情報学の現状と、それを踏まえたさまざまな口承文芸のデータベースの課題へと戻ることとする。

「民謡データベース」は、国立歴史民俗博物館で、館内限定で公開されているデータベースである。データ件数は69000件、文化庁の民謡緊急調査（1979～1989）をもとにしたものである（図1）。基本的には関連するメタデータと、音声データを収録しており、2007年より館内公開を開始した。メタデータ項目は表1のとおりである。

当館内田順子によると、本DBは著作権等については、基本的に研究利用に限り許諾を得ており、データの公開については問題が少ない。原則として、研究者限定の公開であり、その根拠は各都道府県及び文化庁との間で取り決められた「国立歴史民俗博物館 全国『民謡緊急調査』資料活用事業実施要領」によっている。館内のデータベースで必要な情報を検索し、その上で特定の音源等にアクセスしたい場合には、CDにて聞くことが可能である。このCDは各都道府県が作成した録音テープから作成したものであり、これに