

口頭伝承のデータベース化

—その戦略・実践・国際協力の可能性について—

樋 口 淳*

0 はじめに

2015年8月15日のことです。中国社会科学院からメールが届きました。10月10・11日の2日間にわたってCASS Forum (2015・Literature)を開催したいので研究報告をしないか、という話です。テーマは「口頭伝承のデジタル化—その戦略と実践と国際協力の可能性 (Digitizing Oral Tradition: Strategy, Practice, and Collaboration)」で、私が主に日本学術振興会の助成をもとに15年ほど続けてきた「東アジア民話データベース」についての報告を求められているのです。

この種の依頼は、実は3度目です。最初は2010年10月の韓国学中央研究院(ソウル)、つぎが2011年6月の日本口承文芸学会でした。韓国とは2011年の秋にもう一度やり、その時は中国社会科学院も参加しました。

実をいうと私は、こうした話にいささか疲れていました。シンポジウムの招待講演だけではなく、これまでも、機会あるごとに口頭伝承のデジタル保存とアーカイヴ化の必要について説明してきました。そのたびに、誰もが「興味深いアイデアですね。素晴らしいデータベースです」というのですが、その後の手ごたえがないのです。

理由は、おそらく簡単です。国内の博物館の場合は、学芸員や研究スタッ

*専修大学文学部教授

フの専門が分化していて、その専門に従って展示スペースと予算が割り当てられているので、新しい企画が入り込む余地がない。もちろん「企画展」というシステムはありますが、それは一過性の集客力のあるものに限られます。そして「企画展」で大成をおさめ、常設展示のスペースを獲得することは非常に難しいのです。

もちろん、デジタルによる情報発信と、それに連動したデジタル・アーカイヴ、そして展示という連携プロジェクトは、博物館や図書館、そして研究機関にとって重要かつ緊急の課題なのですが、そこでまず必要とされるのは各デジタル・アーカイヴのオリジナリティなのです。「私たちの博物館、図書館、研究機関がいかに素晴らしいデータを保持し（独占し）、いかに素晴らしい独自技術を駆使しているか」が競われるのです。

実はこれは、日本だけでなく、世界中の博物館、図書館、研究機関に見られる傾向だと思います。たとえば韓国学中央研究院の韓国民話と民謡に関するアーカイヴは、素晴らしい出来栄です。なにしろほとんど国家の威信をかけたプロジェクトなので、研究者と技術者が協力し、常設の展示施設をもち、予算規模は民話デジタル化だけで1億円を越え、維持管理費や技術開発は別予算です。しかし、その結果、公開されたアーカイヴは、まさに韓国独自で、韓国世界に閉じ込められ、自己完結的なシステムのもとで運用されることになります。

しかし、デジタル・アーカイヴの次のステップは、言語の制約を越えて、データの互換性を保証し、データを共有して、アーカイヴを利用しながら、アーカイヴを育てる開かれたシステムの構築であることは明らかなです。民俗学や口頭伝承のアーカイヴとは、性格が違いますが、ウィキペディアやユー・チューブは、一つのモデルになるでしょう。

以上のような苦い経験の挙句、さんざん迷いましたが、結局、私は参加することにしました。

会議の共同提案者としてフィンランド文学会の民俗デジタル・アーカイ

ヴ館長のローリ・ハルヴィラーティの名前を見つけたからです。フィンランドは、民話研究発祥の地だし Folklore Fellows Communications という国際通信の拠点です。アメリカからはインディアナ大学のアラン・バーデット、ウィスコンシン大学のロバート・ハワード、ドイツからはロストック大学のクリストフ・シュミットが参加します。アメリカのインディアナは民俗資料収集の拠点として、民俗学研究者なら誰でも知っています。ドイツのシュミットは、ヨーロッパ屈指の民俗資料デジタル・アーカイヴ WossiDIA を主宰しています。もう一人のアメリカ人のロバートは、クラウドを利用した新しい民俗アーカイヴについて語ります。

過去のシンポジウムとは違って、内輪の話にとどまらず、きっと「国際協力」が話題になるに違いありません。それに、私たちの「東アジア民話データベース」の機能性を、おそらく日本やアジアだけでなく、アメリカとヨーロッパに紹介することができそうです。

会議では私が打診したフランス語使用は認められず、おそらくは全中国語と全英語だけの世界になるに違いありません。たった一人の日本人が孤立することは目に見えていましたが、まあ「やってみなはれ」ということでしょう。

以下に「Ⅰ 東アジア民話データベースの試み」として、私が会議の席上で紹介した「東アジア民話データベース」の紹介、「Ⅱ 民話データベースから口頭伝承データベースへ」として、私たちがこれから何を目指しているのかを紹介させていただきます。

Ⅰ 東アジア民話データベースの試み

1. 日本における民話調査・研究の歩み

日本の口頭伝承の記録は、712年に完成した「古事記」からスタートし

たといってよいでしょう。古事記は、太安万侶が稗田阿礼という口頭伝承者の話を文字記録したものとされ、多くの神話や民話が含まれています。

日本では、これ以降、多くの口頭伝承が記録されましたが、このようなタイプの口頭伝承の記録は、時代の差はあれ、世界中に見られると思います。

しかし学術的な口頭伝承記録について言えば、日本においては柳田国男の『遠野物語』（1910）から始まるといってよいと思います。

日本では、『遠野物語』以後、多くの民俗学研究者が、口頭伝承の記録を残してきました。柳田は、『昔話採集の手帖』（1936）という調査マニュアルを日本各地に配布し、地域に残された口頭伝承の調査を促すと同時に、『日本昔話名彙』（1936）というタイプ・インデクスを刊行して、話型による民話の整理に着手しました。

柳田の問題点は、日本固有の民話や民俗の研究にこだわりすぎて、国際比較という視点を軽視したことです。

しかし、柳田の試みは関敬吾によって受け継がれ、日本固有の話型分類とATという国際的なタイプインデクスを連動させた『日本昔話集成』（1950～1958）、その増補版である『日本昔話大成』（1978～1980）という優れたタイプ・インデクスが誕生しました。

1-1. アナログテープによる音声記録

日本の口頭伝承の記録は、テープレコーダーが普及する以前は、主に筆記による文字データだけの記録に止まっていました。

しかし、カセットテープが普及した1960年代後半以後、音声による記録が主流になりました。

記録者は、大学、研究団体、地方自治体、放送局等の機関に属する個人が中心で、記録テープは個人によって保存・管理されているのが一般的です。

その記録数を、正確に把握することは不可能ですが、少なくとも30万件は現存すると思われます。

ただ、これらの記録は、ほとんど全部がアナログ録音であり、時代とともに劣化する危険があります。

1-2. 民話研究・調査の現状

現在でも、日本における民話の記録活動は続いていますし、優秀な話し手もまだ存在しています。

しかし、日本全国に都会化が進行し、民話は、急速に語り手だけではなく、聞き手を失いつつあります。

特に民話の背景となる伝統的な生活様式が失われたために、「民話を語る高齢者の言葉が、話を聞く子供達には理解できない」という現象が頻発しています。

口頭伝承の伝承者が高齢化し、その数は減少しているのに、彼らの話を聞き、次世代に伝える人たちがいません。

その結果、口頭伝承が失われてしまうのです。

正統的な口頭伝承を、どうしたら次世代に継承できるのかが、大きな問題になっています。

2. 東アジア民話データベースの試み

2-1. アナログデータの保存

21世紀を迎え、①アナログの民話研究調査記録の喪失の危機と②語り手と語りの場、そして聞き手の喪失という二つの問題に直面した私たちは、2001年4月に「日本民話データベース作成委員会」を組織し、日本学術振興会の研究助成によって、まず民話研究・調査記録のアナログデータのデジタル保存に着手しました。

すでに述べたように、日本では、oral tradition の伝承者が高齢化してい

るだけではなく、記録者の高齢化も深刻な問題になっています。民話の記録者が高齢化して、時には死亡すると、当該記録者が保有するデータの維持管理が難しくなります。

「データがどこに保管されているのか、どのテープにどの語り手のデータが収録されているのか、何時・何処の調査記録か等の基本情報が消失する」事故が少なくありません。

20世紀の間は、民話記録は文字媒体が主流でしたから、研究者はテープを2次的な資料であると考えて、話を文字化してしまえば、音声資料(audio data)は不要だと考え、音声データを放置するだけでなく、破棄することもよくありました。

もちろん、アナログテープの音声データが、記録者によって、慎重に保存され、カードによって整理されている場合もありますが、その場合にも、困難は残ります。

とくに、日本の場合、民話記録は個人的な性格が強いので、カード整理の方法が記録者によって異なり、記録されたデータの内容が当該記録者にしか分からないことが多いのです。

したがって、調査者が高齢化したり、死亡したりすると、データの整理が難しくなります。私たちが、データの整理とデジタル化を急いだのは、そのためです。

2-2. データベース化の必要

2001年4月から、私たちはまず民話記録者個人に協力を要請し、アナログデータを借りて、デジタル保存し、研究者にテープとデジタルデータを返却し、コピーを委員会に保存するサービスを開始しました。

しかし、この作業の過程で、私たちは、データを単にデジタル保存するだけでなく、検索可能なデータベースにすることの必要に気がつきました。そこで、統一的なマニュアルを作成し、エクセルを使用して、各調査記録

に残された調査者名，話者名，調査日時，話型名，等のデータを項目別に整理する作業を始めたのです。

アナログデータをデジタルに変換し保存するだけでなく，統一基準によって整理し，コンピュータ上にデータアーカイヴを構築することによって，必要なデータを検索し，複数のデータを多岐にわたって比較研究することが，瞬時に可能であることに気がついたのです。

私たちが活動を開始した2001年の時点では，コンピュータのHDDの容量は40G程度でしたから，音声データの保存にはCD-ROMに頼らざるを得ませんでした。

私たちにとって幸いだったのは，その後のコンピュータの進歩です。現在ではコンピュータの内臓HDDは1TBが当たり前です。

音声の圧縮技術も進歩し，MP3を安定的に利用することが出来るようになったので，私たちは，2006年頃から，ファイルメーカーを利用して，さらに機能的なDigital Archiveを作成する作業に着手しました。

現在，私たちが保有する約63500件のデジタルアーカイヴは，64Gの小さなUSBに格納され，どこにでも持ち出すことができます。2001年の時点では，これは夢のような出来事でした。

2-3. メタデータとアプリケーションの問題

2015年現在，東アジア民話データベースのデータは，以下の42の項目に整理されています。

①レコード番号②CD番号③決定題名④話者がつけた題名⑤話者名⑥話者名かな⑦生年月日⑧性別⑨民族名⑩調査時話者住所⑪出身地⑫記録日⑬記録場所⑭探訪者名⑮記録者の所属組織⑯元テープ（オリジナルテープ）番号⑰元テープ管理者⑱分類（民話，伝説，民謡等）⑲大成名⑲大成番号（日本地域カタログ1）⑳AT番号㉑通観番号（日本地域カタログ2）㉒沖縄通観（日本地域カタログ3）㉓伝説大系（日本地域カタログ4）㉔IKEDA（日本地域カタログ5）㉕KT（韓国カタログ）㉖丁番号（中国地域カ

東アジア民話データベース			467 / 63285 no. 1 新規検索 再検索 全データ表示 レイアウト切り替え ? AND/OR追加検索 終了する時はwebページの右上から自分のアカウント名(ゲストアカウント)をクリック。ログアウトをクリックしてください。 クリック検索 ◀ ▶			
レコード番号	CH00E010001	国名: 中国 CH				
CD番号		県名:				
決定題名	オンドリと龍					
話者がつけた題名	ニワトリと龍(鶏和龍)					
話者名	朱蘭(蘭)の母		発音			
話者名かき			結句			
生年月日		性別 女	あいづち			
民族名	漢					
調査時話者住所			伝承事情			
出身地			幼い時に母が語ってくれた			
記録日	1931	写真	文字化資料 1			
記録場所	浙江・紹興		『民間 月刊』1巻8集 P 133-5「鶏和龍」			
採録者名	朱蘭(蘭)		文字化資料 2			
記録者の所属組織			『中国語話集』1(平凡社洋文庫) p23			
元テープ番号			備考			
元テープ管理者			「角」は北京など北方では「チアオ」のように発音するが、南の発音では促音になり、コッコとつづれるの発音に聞こえる。ほかに「雞」と孔雀(よめられるのは龍)、ミスとエビ(よめられるのは目)等多くの類話があ			
分類	11		キーワード			
大成名			鶏 龍 角			
大成番号			梗概			
AT番号	0235		昔ニワトリは角があり、龍にはなかった。ある年、玉皇大帝(天の最高神)が神花を降らせる神を募集したので、龍はニワトリの角を借りて応募する。おかげで龍は雨降に選ばれたが、美しい角を返すのが難しくなり天に送られてしまう。それ以来ニワトリは朝晩天に向かって「龍哥哥(ロンガーガ)、角還我(コックココ)」「龍兄さん、角を返して」と鳴く。			
通称番号	490		よく表示			
冲縄通称						
伝説大系						
IKEDA Ⅱ						
KT						
丁番号	DIN285A動物町鳥(鳥或別動物)	借角或別動物	原話			
金番号	JIN285A公鸡角					
Eberhard番号	EB001オンドリと龍		文獻資料			
話の容量						
公開の了解						
公開の了解者						

图 1

タログ1) ②7金番号(中国地域カタログ2) ②8Eberhard 番号(中国地域カタログ3)
②9話のデータ容量③0公開の了解③1公開の了解者③2発句(語り初めの言葉) ③3結句(語
り終わりの言葉) ③4あいづち(聞き手の相植) ③5伝承事情③6文字化資料1 ③7文字化資
料2 ③8備考③9キーワード④0梗概④1原話(翻訳資料の場合のオリジナル資料) ④2関連文
献資料

以上42項目は、日本の東アジア民話データベースに必要な項目であって、後に検討する国際的な民話データベースには必ずしも必要でないものが多いと思います。(図1)

2-4. 東アジア民話データベース作成上の問題点

私たちの東アジア民話データベースには2つの基本的な問題点があります。

まず第一は、データの入力ミスです。

入力作業は、民話について高度の知識をもった作業員20名ほどが行っています。作業入力には、詳しい統一マニュアルがあります。しかし作業員の能力や関心や判断には違いがあります。

たとえばキーワードの選択には、必ず個人の関心の違いが反映されます。また、その能力や関心や経験や知識の領域にも、必然的に違いが生じます。たとえば、日本の民話についてよく知っていても国際比較についてはよく知らない作業員がいます。逆に国際比較はよく理解できても、日本の民話をよく知らない作業員もいます。韓国民話の専門家には、中国や日本の民話に対する知識が不足しています。

さらに、単純な不注意で入力の手違いをしたり、見落としたりすることも考えられます。実際、東アジア民話データベースには、多くの入力ミスが存在します。

第二は、データの地域的な偏りです。

とくに日本の場合は、民話調査が個人的な努力によって行われましたから、地域的なデータ量の偏りは深刻です。ある地域に熱心な研究者がいれば、当該地域のデータは豊富ですが、その地域に隣接する地域に研究者がいなければ、その地域のデータは空白になります。

もちろん、世界中どこの国の民俗調査でも同じことが、同じことが起こるのですが、日本の場合は深刻です。データの豊富な地域と貧困な地域が混在しています。

2-5. 問題の解決法

この基本的な問題をどう解決したらいいでしょうか。私は、この問題に関してきわめて楽観的です。

まず第一の入力ミスは、気がついた時点で、いつでも訂正できます。データベースのプログラム設計が、しっかりしていれば、誤りを検索することは容易ですから、短時間に簡単に誤りを訂正できます。ファイルメーカーもエクセルも、入力ミスをチェックする上で、たいへん便利なソフトです。

東アジア民話データベースの場合は、資金力と人手が足りませんから、最終責任者の岩倉千春が、たった一人で入力の実の最終チェックをしています。岩倉は、ファイルメーカーによるデータベースの設計者でもあります。もちろん、岩倉のような最終チェッカーがいても、必ずデータの誤りは残ります。しかし優れたチェックのシステムを確立しておけば、誤りはいつでも訂正され、データベースは「より誤りの少ないデータベース」に一步步進化していくでしょう。

つぎに第二の問題ですが、これにも解決策があります。

東アジア民話データベースの場合、これまで主として音声データを入力して来ましたが、これに文字データを加えることで、地域的なデータの偏りは、かなり解消されると思います。

日本の場合には、1910年以來、豊富な民話記録が残されています。民話の調査者は、つねに既存データの少ない地域を目指して調査・研究を重ねてきましたから、過去の文字データが、音声データの欠落を補う可能性は少なくありません。

文字データは、音声データに比べて消失する可能性は少ないので、私たちは、まず消失の危険の高い音声データの救済と整理に注力し、つぎに時間をかけて文字データを入力していけばいいのです。

もちろん、それでもデータの空白地域は残ります。その場合には、根気

よく、補足調査を実施すればよいと思います。

民話の調査には、限りがありませんが、可能なことは全て実行し、データを入力し続ければよいのです。

3. データベースの国際協力と共通プラットフォームの確立と言語

つぎに、私たちの経験に即して、データベースの国際協力と共通プラットフォームの確立と言語、そしてアプリケーションソフトについて考えてみたいと思います。

3-1. 言語の問題

まず、使用言語の問題です。

東アジア民話データベース作成委員会は、2011年から日本・中国・韓国という東アジア三国の民話の国際比較が可能なデータベースの作成に着手しました。

このデータベースは、現在、ほとんどすべてが日本語で出来上がっていますが、将来的には、少なくとも日本語、中国語、韓国語、そして英語の四か国語で検索できるデータベースを目指しています。

この計画は夢のように思えますが、実はそれほど難しいことはありません。それは、コンピュータの技術が、すでに世界化しているからです。コンピュータを利用してデータベースを作成すれば、そのデータベースを多言語化することは難しくありません。

たとえば、私たちが使用しているエクセルやファイルメーカーは多言語対応ですから、入力言語を変えることで基本データがダメージを受けることはありません。

ここで、私たちが崔仁鶴先生の協力を得て、日本語と韓国語使った本当に小さなDBの試作品を紹介したいと思います。

まず、国名選択のプルダウン・メニューから韓国を選択すると、1055件

東アジア民話データベース		1055 / 63285		新規検索		再検索		全データ表示		レイト切替え		?	
レコード番号	KR00C020001	国名:	韓国 KR	NO.	1								
CD番号		県名:											
決定題名	書大栞篇												
話者が付けた題名	書大栞篇												
話者名	黄氏												
話者名かな	みあん し												
生年月日	93歳	性別	女										
民話名													
調査時話者住所	忠清南道善陽郡												
出身地													
記録日	197305	写真											
記録場所													
採訪者名	崔仁鶴												
記録者の所属組織													
元テープ番号													
元テープ管理者													
分類	12												
大成名													
大成番号													
AT番号	0425 0433B 0440												
通観番号													
冲縄通観													
伝説大系													
IKEDA	169 312B 425A 438B												
KT	KT200												
丁番号													
金番号													
Eberhard番号	EB81												
話の容量													
公開の了解													
公開の了解者													

1055 / 63285

NO. 1

新規検索

再検索

全データ表示

レイト切替え

?

AND/OR 追加検索

終了する時はwebページの右上から自分のアカウント名(ゲストアカウントをクリックし、ログアウトをクリックしてください。)

ログイン検索

発句

結句

あいさち

伝承事情

文字化資料

文字化資料 2

備考

キーワード

梗概

原語

文献資料

長く表示

①首女が書大栞息子を産んだ。隣の長者に娘が三人いた。

②書大栞息子が年長し、長者の次娘を嫁にむけた。

③初夜に花嫁は書大栞の夜を拒ぐと美男子が現れた。花嫁は自分の抜けがらを花嫁にあずけて、「誰にも見せてはいけない。病ににおいをさせてはいけない」と念をおした。

④翌朝、嫁が出かけたときに娘たちが書大栞の抜けがらを火鉢にに入れて焼いた。嫁が約束を守らなかったのと一緒にいるわけにはいかないと言いきり、嫁は泣いて逃げた。

⑤嫁は夫に会いたくない一心で翌朝に逃げた。二つの難題を解いて竜宮に向かった。

⑥竜宮には嫁が生んでいて、二人の次女がいた。嫁は娘を見たと、「聞かぬ、美しい女が書大栞息子を産んだ。彼女は書大栞息子を産んだ。中に入れて盗まかせておいた。隣の家は長者の家で、娘が三人いた。ある日、三人の娘が、隣の美しい女が息子を産んだという噂を聞いた。

图 2

のデータ検索が可能であると示されます。(図2)

最初に提示されている話は、崔仁鶴氏が1973年5月に韓国の忠清南道で黄さんという93歳の女性から聞いた異類婚姻譚（人間と蛇が結婚する話）で韓国の分類番号ではKT200、国際比較番号ではAT425, 433, 440に相当する話です。そこで「再検索」してKT200の話を検索すると、1055のデータのうち4件が見つかります。（図3）

国語だけで、中国語や英語のデータは準備されていません。しかし、東アジア民話データベースに中国語や英語のデータを併存させることは、技術的にはきわめて簡単で、特定の言語が他の言語に影響を与えることもありません。

東アジア民話データベースを中国語、韓国語、英語の四か国語による検索可能なデータベースに進化させることは、研究者の協力と、時間と、資金と、人手があれば、いつでも可能だと思います。

3-2. 共通プラットフォームと国際協力

私たちは言語と共通プラットフォームと国際協力の可能性について、次のように考えています。

まず言語については、データベースの作成にあたって、各国の研究者は、自国の言語を使用すればいいと思います。コンピュータは多言語に対応していますから、まったく問題ありません。大切なのは共通プラットフォームです。共通プラットフォームの構築には、国際的な合意が不可欠です。

3-2-1. メタデータの問題

共通プラットフォームを構築する際に、もっとも重要なのは最小限の国際共通メタデータ（入力項目）を決めることです。

入力項目のミニマムを尊重すれば、各国の研究者が、各自のデータベースにどんな入力項目を、いくつ加えるかは自由です。

国際的な共通の項目さえ尊重されていれば、国別・地域別のデータ項目数が国や地域ごとに違ってても、問題は起こりません。

現在のコンピュータの検索能力は、この荷重負担に十分に耐えると思います。

3-2-2. 使用アプリケーションソフトの問題

共通プラットフォームに使用するアプリケーションソフトに関しては、各国のデータの互換性を保証するために、汎用性が高く、しかも比較的安価なソフトを選択するのがよいと思います。

私たちにはファイルメーカーの経験しかないので、「ファイルメーカーは便利だ」としかいいようがありませんが、各国のオリジナルの基本データを、まずエクセルのようなシンプルで、世界中で誰もが使用し、入手可能なソフトを利用して、データの入力と整理をすることが大切だと思います。

それぞれの国の研究者が、それぞれの国や民族の民俗資料をデジタル・アーカイブ化するために、当該資料の性格に即したオリジナル・プログラムを設計することは大切なことですが、民俗資料の国際比較という観点からみると、基本データの互換性のない特殊なデータベースのみを作成しても、国際的にはあまり意味がありません。国際協力を可能にするためには、基本データの互換性が不可欠です。

幸いファイルメーカーは、かなりの互換性がありますので、東アジア民話データベースは、研究者からの要請があれば約63,500件のデータを、無償で提供したいと思います。このデータは64GのUSBに格納できるので、世界中、どこにでもお送りします。

4. インターネット・プラス

つぎに、デジタル・アーカイブが果たす社会貢献（Internet Plus' strategy）についてお話したいと思います。民話のデジタル・アーカイブは、一義的には学術的な研究に役立つことを目指しているのですが、それは研究者以外の一般の人たちにも広く利用される必要があります。

現在、東アジア民話データベースは、ファイルメーカー・サーバーを利用してネット上でデータベース本体を一般公開することを目指して努力を

続いています。しかし63,500件のデータは、量が多すぎて、現在の技術では公開が難しい上に、一般閲覧者の利用には適さないと思います。これは、やはり限られた研究者向けのツールです。

インターネットによる一般閲覧者向けのデータ公開には、利用者の必要性和利用目的にそったきめ細かい配慮が必要です。利用者の必要性和利用目的別にデータを選定し、数を検定して公開する方が効果的です。

私たちのHPサイトは、こうしたアイデアにそって、試行錯誤を繰り返してきました。(http://minwa.fm.senshu-u.ac.jp/EastAsiaDB/IndexEAMinwaDB.html)

現在のところ、私たちのサイトは、残念ながら日本という地域限定です。一応、中国語、韓国語、英語、フランス語、ドイツ語、スペイン語の6か国語に対応していますが、日本語以外のページでは、私たちの活動の紹介とデータのサンプルを提供しているにすぎません。

これに対して、日本語のサイトでは、試験的に幾つかの試みを展開しているの、紹介したいと思います。

4-1. 新しい語り手の登場とその支援

このサイトで、現在、私たちが最も力を注いでいるのは、「世代を越えて継承されてきた口頭伝承を、どうしたら次世代に手渡すことができるか」という問題です。

日本の教育は、現在でも民話の果たす役割を重視し、小学校や幼稚園や図書館や公民館などで、民話を語るイベントがしばしば開催されます。地域の行政機関や観光協会も、地域起しのために、民話を活用したいと考えています。

そして、「子どもや大人に民話を語りたい」と考える人たちもたくさんいます。しかし彼らの多くは、民話を絵本や資料集で読んだ経験はあっても、伝承の民話を聞いた経験がありません。こういう「新しい語り手」を、

私たちはニュータイプと呼びたいと思います。

私たちは、データベースを使ってニュータイプの語り手を励まし、彼らの役に立ちたいと考えています。

4-2. 地域の語りの豊かさを伝える

たとえば、桃太郎のような話を選んで、ネット上で紹介します。

この話は、絵本では、ほとんど一つのヴァージョンしかなく、物語がステレオタイプ化されています。しかし、口頭伝承の世界では、全国各地に様々の桃太郎があります。そして地域の言葉はさまざまで、独特の豊かな表現を持っています。

ですから、たとえば日本の岡山県の子どもたちには、岡山の桃太郎を、岡山の言葉で語ってあげれば、よいと思うのです。地域の教育や観光事業に役立つと思います。

語り手たちが、ともすれば陥りがちな民話のステレオタイプを離れて、地域の言葉で地域の話を書くことのできる、新しい語り手を育てることが、大切だと思うのです。

4-3. 優れた語り手を紹介する

幸いなことに、日本各地には、まだ優れた語り手が存在します。そのために、各地の優れた新しい語り手をビデオで紹介する試みにも着手しています。

日本の新しい語り手たちは、アメリカなどのストーリー・テリングの語り手と同じく、語りに様々の新しい工夫を考えますが、ステレオタイプな語りから抜け出すことができないことがあります。私たちは、彼らに「演出よりも技術よりも、まず伝統的な語りのヴァリエティや言葉の豊かさを身につける方が大切だ」と気づく契機を与えたいのです。

4-4. ユビキタス (Ubiquitous) な資料館

データベースは、このほかにも様々の利用法があるでしょう。インターネットの技術は、ユビキタスです。誰もが、いつでも、どこにでも、手軽に情報を持ち歩き、発信することができます。

ネット上に公開された民話情報を、だれもが共有し、一定のルールのもとに、研究や教育や地域の活性化や観光やビジネスに利用することができます。

地球上、何処にいても、民話と民俗の情報を手に入れ、学び、発信し、新しい可能性を発見することができるのです。

5. 中間的なまとめ

ここで、今回のフォーラムのテーマに沿って中間的なまとめをしたいと思います。

1. 私は、国際的な協力によって、口頭伝承のデジタル・アーカイヴを作成することは難しくないと思います。コンピュータとインターネットの進歩とそのグローバルな能力が、その夢を実現可能なものにしたと考えます。コンピュータは、今後も多言語化に向かって進化し続けるはずです。

2. 国際協力のために、まず大切なのは、メタデータのミニマム・スタンダードを確立することです。国際的なメタデータの数値は、最小限にとどめ、各国・各地域の研究者は、それぞれの国や地域の必要に応じて国別・地域別のメタデータを付け加えればよいと考えます。各国・各地域の自由な探求を保証すべきです。

3. 私がこのメタデータの項目で、必ず必要だと考えるのは、フィンランドのアンチ・アアルネによって確立された話型と話型分類という概念と、現在、国際的に使用されている AT または ATU という話型番号の尊重です。

たしかに AT の分類は、ヨーロッパの口頭伝承の枠組みにそって作成されているので、十分に国際的ではありません。

しかし、これには、デファクトスタンダード (de facto standard) として、すでに100年以上、研究者が利用し、研究成果を蓄積して来た歴史があります。

各国の研究者は、必要に応じて国別・地域別の話型分類を付け加えればいいと思います。

4. つぎに大切な問題は、国際的なデジタル・アーカイブの使用ソフトですが、これは今すぐに一つに決める必要はないと思います。エクセルやファイルメーカーのような汎用性がありシンプルな構造のあるソフトを利用して基本データを整理すれば、将来の国際的なデジタル・アーカイブが、最終的にどのようなソフトを選択しても、問題ないと思います。

5. 最後に小さな問題ですが、音声データの保存に関しても、互換性のあるソフトの使用が大切でしょう。現時点では MP3 が便利ですが、今後のコンピュータの音声圧縮技術の進歩を考慮すれば、基本的な WAVE 形式での保存も検討されるべきだと思います。WAVE はデータ容量が大きいですが、MP3 に比べて互換性が高いというのが、その理由です。現在のコンピュータの能力を勘案すると、WAVE データの管理はむしろかしくありません。

これは、音声データだけではなく、画像データに関しても言えることです。圧縮の技術は日に日に進歩していますが、オリジナルのデータは、大容量であっても、常に圧縮されていないデータを保管すべきだと考えます。

II 民話データベースから口頭伝承データベースへ

0. 東アジア民話データベースの新しい戦略

ここまで、私たちは、民話に特化して東アジア民話データベースの機能や戦略について紹介してきましたが、CASS Forum (2015 Literature) で参加者の研究報告を聞き、討論に参加するうちに2つの大切なことに気がつきました。

第1は、これまで民話に焦点を当てていた東アジア民話データベースを、口頭伝承の全領域を網羅するデータベースに組み替えることです。これは、従来、私たちが見落としていた盲点です。

第2は、東アジア民話データベースの特性を生かし、プログラムを組み替えて、オープンソースとして世界に提供し、口頭伝承の国際比較研究に貢献することです。

1. 東アジア民話データベース・プラス

1-1. 東アジア民話データベースの現在

東アジア民話データベースは、民話の音声記録を重視したデータベースですが、同時に口頭伝承の他のジャンルにも通じたデータベースです。

なぜなら、私たちが民話の調査を行う時には、インフォーマントから彼らの暮らしに関する民俗情報のすべてを聞き取り、暮らしの民俗全体の中に民話を位置づけようとするからです。東アジア民話データベースが、データベース化した約65000件のなかには、神話や伝説や俗信、それに民謡など多岐にわたる情報が収められています。

これまで、東アジア民話データベースはメタデータの中に「分類」という項目を設け、1 昔話(10)、2 伝説(20)、3 世間話(30)、4 神話(40)、5 俗信・ことわざ(50)、6 歌(60)、7 ことばあそび・なぞなぞ(70)、8

東アジア民話データベース			
レコード番号	07M130002	国名: 日本 JP	
CD番号	07M13C001	県名: 福島	
決定題名	会津オオツイ		
話者がつけた題名	会津オオツイ		
話者名	大山ノボヨ		
話者名かな	おおやまのぶよ		
生年月日	19211019	性別	女
民族名			
調査時話者住所	福島県南会津郡舘岩付湯の花		
出身地	福島県南会津郡舘岩付湯の花		
記録日	19960406	写真	無
記録場所	福島県南会津郡舘岩付湯の花		
探訪者名	水谷章三 松谷あよ子 他		
記録者の所属組織	日本民謡の会		
元テープ番号	館岩1		
元テープ管理者	水谷章三		
分類	60		
大成名	無		

2248 / 68285 NO. 329	新規検索 AND/OR 追加検索 クリア検索	再検索 全データ 表示 レイアウト 切り替え ?
終了する時はwebページの右上から自分のアカウント名(ゲストアカウント)をクリックし、ログアウトをクリックしてください。		

発句	無
結句	無
あいづち	無
伝承事情	子どもの頃父親より
文字化資料	無
文字化資料 2	
備考	父親の背中では子守唄代わりに聞いた
キーワード	白虎傳 戊辰戦争
梗概	会津の虎將が滅びる時の歌。

図 4

民俗 (80), 9 その他 (90) という区分をつくり、下位分類してきました。

これを利用して、東アジア民話データベースに収められた民話以外のジャンルを少し検索してみましょう。たとえば、分類の欄に「民謡」のコード60を書き入れて検索すると、2249件のデータがあることが分かります。(図4)

ここに「結婚」というキーワードを入れて検索すると22件の結婚にまつわる歌のデータが検索されます。(図5)

同じように、「俗信とことわざ」のコード50を入力して検索すると1603件のデータが得られます。ここに「雨」というキーワードを入れれば、41件の「雨に関する俗信」のデータが得られます。

以上のように、東アジア民話データベースのデータ検索は、民話を重視していても、潜在的には、民話だけではなく口頭伝承のすべてのジャンルを統合する準備ができています。

東アジア民話データベース		22 / 69285		新規検索		再検索		全データ表示		レイアウト切替え		?	
レコード番号	06S060128	国名:	日本 JP	no.	1								
CD番号	06S06C013	県名:	山形	AND/OR追加検索	終了する時はwebページの右上から自分のアカウント名(ゲストアカウントをクリックし、ログアウトをクリックしてください。								
決定題名	長者の嫁取り	クリック検索											
話者がつけた題名													
話者名	佐藤輝代	発句	むかしむかし すっと昔 あったけど										
話者名かき	さとうてるよ	結句	どんべんからんこ ねっけど										
生年月日		性別	あいづち										
民族名													
調査時話者住所													
出身地		伝承事情											
記録日	19940313	文字化資料											
記録場所	山形県最上郡奥室川町 みるさと伝承館	文字化資料 2											
採訪者名	杉浦邦子 野村敦子 ほか	備考	録音は聴き良い 頃あり 継子譚なし										
記録者の所属組織	昔話の実証と研究 むきのと入 民話と文学の会	キーワード	婆 娘 姥皮 結婚 嫁取り 長者										
元テープ番号	19940313 伝承館A	梗概	長者の家にめい・婆さまが嫁りてきて下働きに雇われる。よく働く婆さまが綺麗な娘になった姿を長者の息子が覗き見る。息子が嫉妬を断る。女中たちが歌い 踊る。婆様が娘の姿について息子の顔つる。										
元テープ管理者	杉浦邦子	表示											
分類	12 60												
大成名	姥皮												
大成番号	209												
AT番号	0510A												

図 5

1-2. 東アジア民話データベースから東アジア民話データベース・プラスへ

そこで、私たちは、2016年度から、民話中心の東アジア民話データベースを修正して、口頭伝承全領域を網羅する東アジア民話データベース・プラス(EAFDBP)に進化させることにしました。

そのために、既存の分類9項目を12項目に拡大し、1 昔話 Folktales, 2 伝説 Legends, 3 神話 Myths, 4 叙事詩 Epic, 5 俗信・ことわざ Superstitions & proverbs, 6 歌 Folksongs, 7 民俗劇 Folk plays, 8 民俗 Folk customs, 9 Folk medicines 民間医療, 10 民間信仰 Folk religions, 11 ことばあそび・なぞなぞ Word play & enigmas, 12 その他 Others とし、その下に以下の表のような120の下位分類欄を用意します。(表1)

2. オープンソース・プログラムの構築

私たちが次に目指しているのは、世界の口頭伝承のすべてを網羅したオ

10. Folktales	20. Legends	30. Myths	40. Epic	50. Superstitions & Proverbs	60. Folksongs
11. Animal tales	21. Religious Legend	31. Origin of Universe	41Yukar	51. Superstitions	61. Work songs
12. Ordinary Folktales	22. Heroes' Legend	32. Origin of Human	42Saga	52. Proverbs	62. Ritual songs
13. Jokes	23. Legend of Origin	33. Origin of Culture	43	53	63. Political satire songs
14. Formula Tales	24. Legend of Another World	34. Origin of Animals & Plants	44	54	64. Daily life songs
15. Erotic Farce	25	35	45	55	65. Love songs
16. Anecdotes	26	36	46	56	66. Children's songs
17	27	37	47	57	67. Historical legend songs
18.	28	38	48	58	68. Others
19.	29	39	49	59	69

70. Folk play	80. Folk customs	90. Folk medicines	100. Folk religions	110. Word play&enigmas	120. Others
71. Puppetry	81. Rite of passage	91. Physical treatment	101. Shamanism	111. Word play	121. Informants' Comments
72. Masque	82. Calendrical and commemorative rites	92. Pharmaceutical treatment	102. Animism	112. enigmas	122
73. Religious play (Mystery)	83. Rites of exchange and communion	93. Religious treatment	103. Fetishism	113	123
74	84. Rites of affliction	94. Magical treatment	104. Worship	114	124

74	85. Rites of feasting, fasting and festivals	95	105	115	125
76	86. Political rituals	96	106	116	126
77	86. Others	97	107	117	127
78	88	98	108	118	128
79	89	99	109	119	129

表 1

ープンソース・プログラムの構築です。

CASS Forum (2015 Literature)に参加した中国、フィンランド、ドイツ、アメリカ、日本の研究者の誰もが認めたように、口頭伝承資料のデジタル保存とアーカイブ化は、世界の趨勢です。そして世界各地の民俗資料館は、それぞれが所蔵する貴重な資料のデータベース化を推進しています。

私は、この会議での討議を聞きながら、そこには2つの大切な問題があり、その2つの問題は互いに矛盾する違った方向に向かうかもしれないと考えたのです。

2-1 大切な2つの問題とその矛盾

今回の CASS Forum (2015 Literature) のメインテーマは、口頭伝承資料のデジタル化の戦略と実践と国際協力の可能性でした。そこで話し合われたのは、①デジタル化の実践に際して、世界各国・各民族が民族資料の独自性をいかに尊重するかということと、②国境や民族の枠を超えて、いかにネットワークを構築し、各地の民俗遺産を人類共通の遺産にするのか、という2つだったと思います。

口頭伝承の研究者なら誰でも気がついているように、口頭伝承は地域や民族に固有のものでありながら、同時に世界に向かって開かれています。たとえば「シンデレラ」は、中国にも、フィンランドにも、ドイツにも、

アメリカにも、日本にもありますが、地域ごとに暮らしに根差した独自性を有しています。

ところが、世界各地の口頭伝承資料館が、地域独自の文化とライフスタイルの特色を生かし、所蔵資料の独自性に焦点をあててデジタル資料を提示しようと努力しようとすればするほど、プログラムは独自のものとなり、資料の互換性を失うことになります。

たとえば、東アジア民話データベースのデータの中の沖縄県石垣島で記録された民謡について考えてみましょう。その民謡は石垣島方言で歌われていますが、石垣島方言は、独特で地域の人以外には理解が難しいです。そこで、多くの人にアクセスしてもらうためには、音声データに文字データを加える必要があります。しかし石垣島方言の音韻体系は、共通日本語の音韻体系とは違いますから、できるだけ忠実に表現するためには特別な工夫が必要になります。さらに、歌の理解を深めるためには、歌の音データの進行に合わせて、歌われている箇所の文字データを歌の進行に従ってたどるマーキングも必要になります。

このような複雑な操作は、現在のコンピュータの能力とプログラミングの技術を考え併せれば、それほど難しいことはありませんし、世界中の多くの民俗資料館が既に採用している一般的な手法です。日本の民謡データベースで、現在、このように行き届いたデジタル化が行われているかどうかは寡聞にして把握していませんが、将来はおそらくそうなるだろうと推測します。

しかし、ここで問題になるのは、データの互換性と比較研究上の障害です。

石垣島民謡の独自性を生かしたデジタルデータは、たとえば沖縄全域の民謡データや、日本全国の民謡データと統合して、「沖縄民謡アーカイヴ」や「日本民謡アーカイヴ」を構築することが可能かもしれませんが、これをフィンランドやドイツやアメリカの民謡アーカイヴと合体させて「五か

国民謡アーカイヴ」を構築することは、現状では難しいし、もし出来ても、あまり意味がないと思います。

しかし、口頭伝承の研究の上では、例えば「結婚」というキーワードでフィンランド、ドイツ、アメリカ、日本の五か国の民謡を検索したり、「結婚」に関する「習俗」と「昔話」と「俗信」と「歌」をクロス検索して、比較研究することが、大変重要です。

地域の独自性や民謡というジャンルに特化したプログラムは、残念ながら、こうした国際比較のためには十分な力を発揮することができません。とくに口頭伝承のジャンルを横断したクロス検索には、対応が難しいはずです。

2-2. 矛盾の克服

世界のデジタル・アーカイヴが、膨大な予算を傾けて、オリジナルなプログラムを開発し、独自の資料公開を目指している現在、この2つの問題の矛盾は深刻で、克服しがたいようにみえます。もちろん、コンピュータの技術や能力が飛躍的に拡大し、クラウドが大容量のデータ共有スペースを提供するに違いない近未来には、何が起こるか分かりません。今日、解決が難しいと考えられる問題も、すべて氷解する可能性はあります。

しかし、私は、今すぐにでも、この矛盾を解決する方法があると考えています。

それは、まさに国際協力によって、東アジア民話データベースのようなシンプルで互換性のあるプログラムをオープンソースとして共有し、各自のデータを共有すればいいのです。

そして、同時に、世界各地の研究者は、各自の地域の民俗資料のデータの独自性に配慮した独自のプログラムに基づくデジタル・アーカイヴの構築に励めばいいのです。

世界各地の独自性にあふれたデジタル・アーカイヴの中の片隅の一室に、

オープンソースの共有資料検索のコーナーを設置し、研究者であれば誰でも、目的に従って世界各地の民俗資料と、自らの地域の資料をクロス検索できるようなサービスを提供すればいいのです。

USB や、小型の HDD に用意に収納されてしまう東アジア民話データベースプラスのようなデータベースは、この役割を果たすのにふさわしいと思います。

私たちは、つぎにオープンソースの一つのモデルを提案したいと思います。

2-3. オープンソース・プログラムに関する一つの提案

すでにお話しした通り現行の東アジア民話データベースは、エクセルとファイルメーカーというシンプルで汎用性の高いソフトを利用しているので、ユニコードに対応していれば、世界中のほとんどの言語を入力することができます。

問題となるのは、現在42もある日本独自のメタデータです。
そこで、これをたとえば

- ①国別データ番号 national data number ②データの名前 title ③国名 nation (←pulldown menu) ④民族名 ethnic group ⑤分類 classification (←pulldown menu) ⑥分類コード classification code ⑦インフォーマント名 name of teller/informant ⑧性別 Sex (male/ female) ⑨生年月日 date of birth ⑩住所 address of informant ⑪出生地 place of birth ⑫調査年月日 date of research ⑬調査地 place of research ⑭調査者名 name of researcher ⑮所属組織 organization ⑯国際カタログ番号 international catalogue number AT/ATU ⑰伝承事情 Informants' Comments & notes on transmission ⑱キーワード key words ⑲文字化された梗概または全データ summary or whole original written data ⑳備考 comments ㉑公

開許可の有無 permission for open access ②公開許可者名 permission given by ③参考文献 references

という23にしぼって、世界中のどの地域のデータも受け入れることのできるメタデータ・スタンダードを作成してみたいと思います。

つぎに問題になるのは、口頭伝承の分類です。

東アジア民話データベースは東アジア民話データベースプラスに進化する過程で、既存の分類の9項目を12項目に拡大し、

1 昔話 Folktales, 2 伝説 Legends, 3 神話 Myths, 4 叙事詩 Epic, 5 俗信・ことわざ Superstitions & proverbs, 6 歌 Folksongs, 7 民俗劇 Folk plays, 8 民俗 Folk customs, 9 Folk medicines 民間医療, 10 民間信仰 Folk religions, 11 ことばあそび・なぞなぞ Word play&enigmas, 12 その他 Others を採用することにしました。(表1)

この12の分類は、メタデータ⑤の分類 (classification) の中に pulldown menu として組み込まれています。

口頭伝承の分類が、果たしてこの12項目でよいのかどうかは、大きな問題ですが、さらに複雑なのは、その下位分類です。

今回は、一つの試みとして12の分類項目のもとにそれぞれ10の下位分類欄を用意し、たとえば「歌」の下位分類には、中国の呉超氏の分類に従って

①労働歌 Work songs -61 ②儀式歌 Ritual songs -62 ③時政歌 Political satire songs-63 ④生活歌 Daily life songs-64 ⑤情歌 Love songs-65 ⑥児歌 Children's songs -66 ⑦歴史伝説歌 Historical legend songs -67 ⑧ 其他 Others

を用意してみました (cf. 『中国民歌』 呉超著・浙江省教育出版社・一九九六年三月)。

呉超氏は、たとえば児歌を

A 揺籃曲（眠らせ歌）、B 数数歌、C 問答歌、D 遊戯歌、E 連鎖調の5つに下位分類していますから、そこに

A 揺籃曲（眠らせ歌）66a、B 数数歌66b、C 問答歌66c、D 遊戯歌66d、E 連鎖調66e

のような下位分類を付け加えることも可能だと思います。（表1）

しかし、世界各地の「歌」には、それぞれ独自の分類案があるはずですから、慎重に検討する必要があります。

メタデータのミニマム・スタンダードと口頭伝承の分類は、一度スタートしてしまうと簡単に変更することはできません。ですからスタート前によく考える必要があるのです。

私たちは、口頭伝承の分類について考えるとき、いつもフィンランドのアアルネとアメリカのトンプソンの協力を思い出します。アアルネは1910年に刊行されたFFC3号に最初の『昔話の話型カタログ *VERZEICHNIS der MÄRCHENTYPEN*』を発表した時、すでに2000の話型を用意していました。このカタログは、わずか63ページでした。そしてトンプソンがこのカタログを増補して1973年に刊行した刊行したFFC164号の『昔話のカタログ *THE TYPES of FOLKTALE*』は588ページで、ページ数は10倍ほどに増加しました。しかしトンプソンは、そこに形式譚と分類できない話という新しい話型500を付け加えただけで、アアルネの最初の分類を尊重したのです。

アメリカ先住民の口頭伝承資料に深い理解を持ちながら、民話の国際比較というアアルネのアイデアを推進した彼の努力を、私は高く評価したいと思います。

民話の研究者なら誰でも知っているように、アアルネのカタログはヨーロッパの民話中心で、アメリカ先住民や中国や日本の民話の分類には適していません。そこで、各地域の研究者はアアルネの考えた話型というアイデアを尊重しながら、各地域独自のカタログを作り、アアルネのアイデア

を利用して、民話の国際比較研究をしているのだと思います。

私は、口頭伝承のデジタル・アーカイヴを作成し、国際比較研究に役立てることは、非常に意義深いことだと考えています。

このデジタル・アーカイヴは、アアルネとトンプソンのカタログという素晴らしい遺産を包括的に受け継ぐだけではなく、これまでに展開されたすべての国際比較研究の業績を包み込み、さらには口頭伝承のすべてのジャンルを網羅して、比較検討する手段を与えるという可能性があります。

アアルネとトンプソンのカタログの場合と同じく、このオープンソースのメタデータや分類の構成が、すべての研究者を満足させることは難しいでしょう。しかし1910年にアールネが初めてのカタログをスタートさせたように、私たちもできるだけ早い機会に、このオープンソース・プログラムを発足させたいと願っています。