

民話データベースの課題と可能性

樋口 淳

1 日本の民話記録とデータベース化作業

1-1. 日本の民話記録

二〇一〇年は、柳田國男が『遠野物語』を著してから百年にあたります。日本では、『遠野物語』以前にも、民話の記録は、さまざまな形で残されていますが、「聞き手が特定の語り手から語りを聞き取り、意識的に記録した」という意味では、やはり『遠野物語』は一つの時代を画し、日本の口承文芸学研究の第一歩を記したと言つてよいでしょう。

これ以降、次第に緻密化されていく民俗学・口承文芸学の枠組みの中で、多くの研究者が、優れた語りの記録を残してきました。こうした事情は、ヨーロッパでも同じで、一八一二年の『グリム童話集』の刊行を一つの画期とすれば、日本よりも百年ほど早く記録が始まり、十九世紀の末には極めて質の高い資料集が各地で生まれていたと言つてよいでしょう。

しかし、こうした記録の作業は一九六〇年代に入ると一変し

たと言つてよいと思います。テープレコーダの普及とともに各地で音声をとまなう忠実な記録活動が始まったのです。とくに日本ではオープンリールに続いてコンパクトなカセットレコーダが普及し、誰もが簡単な操作で音声記録を残せる環境が整いました。

日本の場合は、テープレコーダによる記録の活用が特に盛んで、二十世紀の終わり頃までには二〇万件とも三〇万件とも考えられる世界有数の音声記録が残されました。

1-2. 民話記録の保存とデジタル化の必要性

民話の記録に携わるものなら誰もが気がついているように、二十世紀の終焉とともに、伝統的な語りの場に大きな変化が訪れました。もちろん現在でも、村々を訪れて語りを記録している方々は少なくありませんが、その一方で一九八〇年代代までは気楽に聞くことのできた話が、つぎつぎと消えていったことも事実です。それと同時に、一九六〇年代から記録を続けてきた

話の記録者・研究者の高齢化も進行しました。

語りの音声記録の方法には、テープやカード整理など一定の共通ルールはありますが、記録者ごとの手づくり部分も多く、記録の大半は記録者が作業を放棄すると、ただの粗大ごみになってしまいます。幸運にも地域図書館や大学研究室に残された場合でも、管理者が変わると素性の分からない「使えない」資料になって埃をかぶってしまうのが実情です。

私たちは、二〇〇一年四月に「日本民話データベース作成委員会」、二〇〇六年四月に「沖縄伝承話データベース作成委員会」という二つの組織をスタートさせ、全国の研究者から時とともに劣化するアナログテープの記録資料を借り受け、デジタル化し、集中管理することを目指しました。デジタルの資料は、操作を誤ると一瞬のうちに消失してしまうという恐ろしい危険もあります。複数の担当者が、複数の手段で保存すれば、恒久的に保存することが可能です。

私たちが作業をはじめた二〇〇一年には、コンピュータの能力も低く、音声処理もままなりませんでしたが、現在ではきわめて簡単な作業になりました。デジタル化した資料を保存する手段も、飛躍的に向上し、外付けのディスクを利用すればテラバイト単位のデータ保存も難しくありません。

さらに音声データの圧縮技術も進歩し、MP3などの普及によって、たとえばmp3のような小型ツールのなかに二万話以上のデータを収納できるようになったのです。

1-3. データベース化のメリット

こうしたデジタル技術の進化と信頼性の向上が進む一方で、次第に明らかになってきたのが、データベース化のメリットです。従来の民話のアナログ記録には、実は大きな問題がありました。

たとえば「沖縄伝承話データベース」の場合、その基礎となった沖縄国際大学の遠藤庄治さんと学生達の調査記録は、これまでの日本の民話記録のうちで最も体系的なものです。この記録はカセットテープ約四、五〇〇本に収められ、ほとんど全ての資料がカードに整理されています。この資料は、現在データベース化作業の最中ですが、作業が完了すること、テープは沖縄県立博物館の収蔵庫に保管されて行きます。日本の民話記録のうちで、このような手順で保存された記録は質量ともに他にないと思います。

しかし、誰かがこのアナログ資料を利用して、たとえば「アカマタ婿入り」と「浜おり」の関連について研究しようと思いついたとしましょう。その研究者は、何日もおもろ町の沖縄県立博物館に通いつめて、テープカードを調べて沖縄各地の「アカマタ婿入り」と「浜おり」についての記録を確認し、当該テープを見つけ出し、テープレコーダにかけ、早送りとまき戻しを繰り返して「アカマタ婿入り」と「浜おり」の語りを探し出さなければならぬのです。この作業には、膨大なコストがかかります。よほどの覚悟がないかぎり不可能な仕事です。

これに対して、「沖縄伝承話データベース」が完成すれば、小さなノートコンピュータに全ての資料が収納されているわけですから、東京にいやと北海道にいやと、たとえば「キーワード検索」で「アカマタ」「浜おり」と入力することで、必要なデータのほとんどを瞬時に入手することができます。

2 日本民話データベースと

沖縄伝承話データベースの現状と課題

2-1. データベース化作業の現状

こうして「日本民話データベース作成委員会」は、二〇〇一年四月に作業を開始し、以後五年間にわたって日本学術振興会から助成を受けて、二一、〇〇〇件の民話に関するデータを処理し、データベース化しました。

学術振興会による助成は大体五年が上限ですから、私たちは二〇〇六年度は一度活動を休止し、その間にデータのプログラム・ミスなどをチェックしました。

(二〇〇五年度の研究助成終了時点での日本民話データベース作成委員会のメンバーは、岩倉千春、小野和子、剣持弘子、高津美保子、武田正、竹原威滋、立石憲利、田畑千秋、常光徹、花部英雄、牧夕野靖子、丸山顕徳、三原幸久、山下欣一、米屋陽一です。)

そして二〇〇七年四月からは、沖縄伝承話資料センターと協力して「沖縄伝承話データベース作成委員会」を立ち上げ、地域を沖縄に限って民話関連データの処理を行うこととし、最終

的に二七、二〇二件を処理する予定です。

「日本民話データベース」と「沖縄伝承話データベース」は、それぞれ独立していますが、まったく同じプラットフォームの上で作業が進められているので、現在は「日本民話データベース」「沖縄伝承話データベース」「統合版・日本民話データベース」の3つが機能し、二〇一二年四月現在には「統合版・日本民話データベース」を使用すれば五万件程度の日本民話関連資料を検索できます。

2-2. データベース化の技術的な問題点とその克服

この間、私たちが日本民話データベースに着手した時に直面した大きな問題は、二つありました。

2-2-1. データベースの公共性と個人の業績の保証

一つは民話記録のデータベース化ということに対する理解の不足です。すでに述べたとおり、日本民話の音声記録は一九六〇年代に始まり、以後四十年の間に、少なくとも約三〇万件程度のデータが記録されたと推測されますが、それらほとんど全ては研究者個人の努力によるもので、記録は個人によって所有されています。研究者個人は、自分の調査に個人的な愛着がありますから、日本民話データベースのような公共的性格の強い資料体に、情報を提供することに消極的です。

ですから、最初は記録タイプの借り出しに時間がかかりまし

た。日本民話データベース作成委員会は、当然のことながら、記録テープをデータ化する際に「当該テープの所有者が誰か」「調査した組織は何か」「それぞれの話を聞いた調査者は誰か」を明らかにし、データベース化された資料から、テープ所有者、調査組織、調査者等の名前が検索可能なものとし、それぞれの研究者の個人的な業績を保証しました。その上で、アナログのテープを借り出して、データベース化の作業を行い、作業が終わると、CD-ROMにデータを記録し、資料提供者にテープとともに返還しました。研究者の中には、コンピュータを利用しない人もいたので、オーディオCDも用意して、一般のステレオでも簡単に記録を確認できるようなサービスを行いました。

資料を所有する研究者たちは「今後、自分の調査した資料が、どう使われるか」に対して不安を抱いていましたから、次第に資料提供に対して肯定的になり、最終的には限られた予算では処理しきれないほど大量のデータが集まりました。

2-2-2. データベース入力作業者の養成とその業績の評価

二番目に直面した大きな問題は、作業者の確保です。アナログテープをデジタルデータにするのは、簡単な機械的作業ではありません。極めて熟練のいる作業です。

まずアナログ録音のノイズを取り除きながらデジタルに変換する作業があります。これは比較的簡単な機械的な作業で、誰でもできます。

次に、このデータを一件ごとに切り分ける作業があります。これは、調査者によって作業手順が違うので、個別の対応が必要で、深い経験が必要とする、複雑な作業です。

この場合、一番簡単なのは、調査者自身が「これから何処で、誰の、何という話を聴きます。それでは始めます」といって調査を始め、「これで記録を終わります」といってテープを止める調査です。この場合は、調査者自身が一件毎にデータを整理して録音しているので、作業は簡単です。テープが貴重であった初期の調査では、この方法が多く採用されました。

しかし、次第に録音技術も進歩し、テープが安く大量に入手できるようになると、一時間半あるいは二時間の間テープが途切れなく録音されることが多くなります。民話の記録調査を経験した人なら誰でも理解できるように、一つの話が語られる合間の雑談の時間に、語り手が当該の話に関して貴重な情報を提供するものが少なくないからです。調査の時には、いつ貴重な民俗情報が提供されるか予測できないわけですから、テープは調査の間中、途切れなく記録しておくのがよいのです。

こういう調査全体を記録したテープを整理するのは大変な作業です。調査者自身には、どこに重要な情報があるのかが分かっているからよいのですが、第三者が情報を整理するためには、まず全体を聴き、テープの何処に重要な情報があるかを判断し、テープの情報を一件ずつ切り分けなければなりません。

したがって、テープ整理は「民話に対して深い理解と愛着」

のある人でないとできない仕事になります。民話をよく知っていないと、何が大切な情報であるかを理解できないので、機械的に大切な情報を切り捨ててしまうことが起こります。テープの整理には、一件ごとに三十項目ほどのデータを入力するという細かい作業が不可欠ですが、こうした細かい入力作業よりも、話を一件ごとの情報単位として切り分ける作業が重要です。

作業者は、最初はコンピュータの使用法から学び始め、テープの整理をするうちに少しずつ「大切な情報が何であるか、何処にあるか」を学んでいきます。「日本民話データベース」の場合は、だいたい日本民話の会に所属する女性が作業を行い、「沖縄伝承話データベース」の場合には、かつて沖縄国際大学で学び、民話の調査を経験した人達が、主として自分の調査したテープをデータベース化しています。いずれの場合も、コンピュータの作業手順を理解した後は、非常に質の高い仕事をしています。

3 データベースの紹介

3-1. 作成ソフト…ファイルメーカー

それでは次に「実際のデータベースで何ができるか」を具体的に紹介したいと思います。

データベースは、日本民話データベースも沖縄伝承話データベースもファイルメーカーというソフトを利用して作成されています。作成担当者は、両データベースの作成委員会委員の岩倉千春さんです。

ファイルメーカーは、ウインドウズとマキントッシュの二つのOSでプログラムを作成できるソフトですが、完全なプログラムはマキントッシュで組まれています。

音声データは、基本的には π 形式で保存されていますが、ノートパソコンやPDAを利用する場合には、MP3の形式で保存されたデータが利用されています。

3-2. 検索項目33

さて、このデータベースの検索項目は、次の三十三項目です。

- 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調査時話者住所、⑩出身地、⑪記録日、⑫記録場所、⑬探訪者名、⑭記録者の所属組織、⑮元テープ番号、⑯元テープ管理者、⑰分類、⑱大成名、⑲大成番号、⑳AT番号、㉑（通観番号）、㉒発句、㉓結句、㉔あいづち、㉕伝承事情、㉖文字化資料、㉗備考、㉘写真、㉙キーワード、㉚梗概、㉛話の容量、㉜公開の了解、㉝公開の了解者

このデータベースを開くと図1のようなページが開かれ、番号は記されていませんが、「語り手の名前」や「話の題名」など、項目別に記入事項が書かれています。たとえば、上の図の 경우는、語り手は山形県米沢市の「平田幸一」さんで、話の題名は「桃太郎」、記録したのは「水野道子」さんで、一九八四年八月のことだったなどということが分かります。

日本民話データベース		2 / 21437 no. 1		新規検索	再検索	新規ウインドウ	とじる	?
レコード番号	06M160007	県名:	山形	AND/OR	追加検索	レアウト	全データ	印刷
CD番号	06M160001			クリア検索	ホームページ	切り替え	表示	
決定題名	桃太郎			クソの検索	録り	me		
話者がつづけた題名	桃太郎					ライブラリ		
話者名	平田肇一	伝承事情	読本					
話者名かひ	ひめたごういち	文字化資料	不明					
生年月日	19110904	備考	楽団子の部分なし					
調査時話者住所	山形県米沢市塩井町	キーワード	桃 鬼 犬 キジ					
出身地	山形県米沢市塩井町							
記録日	19840800	複製	無					
記録場所	山形県米沢市							
探訪者名	水野道子	検索	桃太郎が大きくなって鬼退治に行く。犬やキジが家来になる。鬼を退治して帰る。					
記録者の所属組織	小平民話の会	長く表示						
元テープ番号	水野1							
元テープ管理者	水野							
分類	12							
大成名	桃の子次郎							
大成番号	143							
AT番号	513A							
通称番号	不明							
伝説大系								
発句	無							
結句	無							
あいづち	無							
		話の容量	22.761KB	2分12秒				
		公開の了解	不明	公開の了解者:				

これを新しく検索しなおして、たとえば「岡山」県の「桃太郎」について知りたければ、図2の決定題名に「桃太郎」、記録場所に「岡山」と書き込んで「実行」のボタンをクリックすればよいのです。(図1)

日本民話データベース		検索条件入力	
レコード番号		県名	
CD番号			
決定題名			
話者がつづけた題名			
話者名			
話者名かひ		性別	
生年月日		写真	
調査時話者住所		実行	
出身地		キャンセル	
記録日			
記録場所			
探訪者名			
記録者の所属組織			
元テープ番号			
元テープ管理者			
分類			
大成名			
大成番号			
AT番号			
通称番号			
伝説大系			
発句			
結句			
あいづち			
伝承事情			
文字化資料			
備考			
キーワード			
梗概			

4 入力作業の実際

検索は、いくらでも複雑に行うことができます。たとえばこれに>という記号と一九七〇年とを使用し「V1970」という条件を加えれば、一九七〇年代以前の調査で記録された「岡山」の「桃太郎」などというデータも抽出できます。キーワードに狐や虎を入れれば、話型に関わらず、虎や狐が登場する話のデータも抽出できます。梗概に虎や狐を入れれば、キーワードとして採用されなかった虎や狐の登場する話のデータをクリッカー一つで入手できるのです。(図2)

次に、データベース作成の作業手順についてお話します。

4-1. アナログデータのデジタル化
 作業者は、まずアナログテープをWave形式でデジタル化します。変換ソフトは、なにを

利用してもかまいませんが、テープによってはノイズが大きいのので、ノイズを除去する作業が必要になります。調査中にテレビの音や電話の音が録音されてしまうこともあります。あまり神経質にノイズを除去する必要はないと思います。

これに対して、現時点では音声をWave形式で保存することが、絶対に必要です。Wave形式で保存しておけば、これをMP3あるいはMP4に変換することは簡単です。オーディオ・ファイルに変換して、コンピュータ以外の一般のオーディオ・プレーヤーで聴くこともできます。まず、この作業を行い、オリジナルのテープをWaveに変換して、ハード・ディスクに保存すると同時に、CD-ROMにもコピーして保管します。

4-2. データの切り分け

作業者は、次にオリジナルのテープに収められていた九〇分あるいは一二〇分の情報を、話ごとに、一件ずつ切り分けます。この作業は、たいへん熟練を必要とします。未熟な作業者は、何が大切な情報かを判断できず、機械的に作業を進め、重要な情報を削除してしまう可能性があるからです。

また、たとえ作業者が熟練していても、調査時の記録が不十分であったり、ノイズが大きくて、十分に聞き取れなかったりして、当該の記録を保存すべきか否か、判断に迷う場合もあります。日本民話データベースと沖繩伝承話データベースの場合は、不十分と思われるデータも、必ず一件として保存することにしました。

とくに、話者が話を途中で忘れてしまったり、語り間違えたり、訂正したり、データとして不完全な場合は、必ず保存するようにしました。

以上のように、データの切り分け作業には、さまざまの問題が発生します。したがって、多くの作業者が関わる場合には、最初によく話し合い、作業に共通理解を徹底することが必要です。そして、作業者が判断に迷った時には、熟練した作業者に相談できるような連絡体制を作る必要があります。

4-3. エクセル入力

テープの切り分けが出来たら、つぎはエクセルを利用して、六ページで紹介した三十三項目の情報を入力します。

図3の場合は、エクセルの空欄に、全ての項目が入力されていますが、これは例外です。資料によっては、話者名も調査者名も、調査地や調査日時も不明なものがあります。しかし、資料はどんな資料でも大切です。それに、内容さえ確認できれば分類は可能ですし、使用された方言などから地域を特定できるかもしれません。ですから、エクセルの入力には、空欄があってもかまいません。(図3)

4-4. 完成データの提出とデータベースの作成

以上の作業を終えた作業者は、①オリジナルテープ、②オリジナルテープをデジタルに変換したCD-ROM (Wave) 一枚、

	A	B	C
1	レコード番号	040020221	040020222
2	CD番号	040020019	040020019
3	決定題名	金の弁	ネズミづなつた姉娘
4	話者がつけた題名	金の弁の話	ネズミづなつた娘
5	話者名	永浦誠喜	永浦誠喜
6	話者名かな	なからせいき	なからせいき
7	生年月日	19090417	19090417
8	性別	男	男
9	調査時話者住所	宮城県登米郡南方町青島屋敷72	宮城県登米郡南方町青島屋敷72
10	出身地	宮城県登米郡南方町青島屋敷72	宮城県登米郡南方町青島屋敷72
11	記録日	20000130	20000130
12	記録場所	宮城県登米郡南方町ハの森40-1 南方町公民館	宮城県登米郡南方町ハの森40-1 南方町公民館
13	採訪者名	小野和子・山田和郎・山田裕子・秋山伸司	小野和子・山田和郎・山田裕子・秋山伸司
14	記録者の所属組織	みやぎ民話の会	みやぎ民話の会
15	元テープ番号	67A	67A
16	元テープ管理者	小野和子	小野和子
17	分類	12	12
18	犬成名	黒金の弁	灰坊 笛吹き髯
19	犬成番号	226	211 119
20	AT番号	AT729	AT510
21	〈通称番号〉	52金の弁	181灰坊 223月女屠
22	発句	わかえい	わかえい
23	結句	無	無
24	あいづち	うん んで	うん んで
25	伝承事情	祖母よふさんから、田伊寒嶋・かまどの焚き口などに座って姉とよく聞く。	祖母よふさんから、田伊寒嶋・かまどの焚き口などに座って
26	文字化資料	永浦誠喜語り みやぎ民話の会編みやぎ民話の会叢書第十集 青島屋敷老翁夜話 甲巻Ⅱみやぎ民話の会 2001p.66-67	永浦誠喜語り みやぎ民話の会編みやぎ民話の会叢書第十集 巻Ⅱみやぎ民話の会 2001p.68-73
27	備考	話者永浦誠喜さんは2002年9月7日逝去。誠喜さんは昔話を定型句で語り納めるのを好むが、ここでは会話の流れで結句を欠いている。	話者永浦誠喜さんは2002年9月7日逝去。誠喜さんは昔話を定型句で語り納めるのを好むが、ここでは会話の流れで結句を欠いている。
28	写真	有	有
	キーワード	木(仮)の翁 湯重 金の弁と金の弁 コイ 隣の爺 火くわい	長崎徳右衛門 長(徳(な)か) 村祭り 能楽奉納 笛吹き

図 3

③ データごとに切り分けた「完成CD-ROM」(Wave)二枚、
 ④ エクセルに整理されたデータ、という四つをセットにして管理責任者に提出します。

これを受け取った管理責任者は、ファイルメーカー等のソフトを利用して、データベースを作成します。

5 日本民話データベースの問題点

5-1. 地域的な偏り

すでに述べたように、日本における民話調査は、そのほとんど全てが個人の研究者によるものです。したがって、地域による偏りが多いのです。たとえば、日本民話データベース作成委員会メンバーの立石憲利さんは、きわめて優れた民話研究者で五十年にわたって調査を続け、多くの民話資料集を公表しています。彼のフィールドは岡山県ですから、岡山周辺の資料は実に豊富です。また、宮城県の小野和子さんは、宮城民話の会を組織して、仙台を中心に現在も調査活動を行っています。さらに沖縄の遠藤庄治さんが沖縄全域の調査を実施し、すぐれた資料を残したことは、すでに述べた通りです。しかし、これらは例外で、たとえば近畿地方などは、ほとんど手つかずの状態です。データベースには空白地域が多いのです。

しかも、民俗の分布は、複雑で、山や川を一つ隔てれば、まったく言葉や習慣が違うことが一般です。Aという地域を集中的に調査しても、すぐ隣のBという地域には、まったく違った民

俗がある可能性があります。したがって、どれほど調査しても十分ということはないのですが、日本民話データベースの場合には「基本的な資料に地域的な偏りがあるという欠陥がある」という批判は免れえません。

5-2. 作業者による入力作業の偏り

これも既に述べたことですが、複数の作業者が「データの切り分け」と「エクセル入力（情報入力）」に携わります。そこには、当然のことながら個人の能力や嗜好の差が表れます。たとえば、「キーワード」として何を選ぶか、「梗概（要約）」にどんな基準を設けるか、などには、個人差が生まれます。基本的な「話型」の決定に関してすら、作業者の能力や個人的な判断が影響します。

したがって最終的に完成したデータベースにも、多くのエラーが発生します。データベースのプログラムを作成し、作業者の作成したエクセルと音声のデータを次々と加えていく「プログラム作成責任者」は、当然、その作成作業で個々のデータの不備や不足をチェックして、自ら修正したり、作業者にフィードバックして作業をやり直させたりするのですが、それにも限界があります。

データベース作成の作業は、膨大な労力を必要とする作業ですから、ある場合には「ある程度の不足には目をつむる」ことが必要な場合もあります。

6 データベースの問題点を、どう解決するか

6-1. 地域的な偏りを、どう補うか

日本民話データベースの地域的なデータの偏りは致命的なものです。今後、学術振興会の援助をどこまで得られるか明らかではありませんが、まず為すべきことは、資料空白地帯において行われた過去の調査資料（アナログテープ）を借り受けて、データベース化を推進することです。日本民話データベースがまだ処理していないアナログ資料は、少なくとも十万件はあると考えられます。沖縄だけでも四万件の処理が残されています。

しかし、次にまず着手すべきことは、音声資料だけではなく過去の文字資料をPDF等のファイルにして、統合版日本民話データベースに付け加えることだと思われれます。文字化された資料は、音声資料よりも長期にわたって保存が可能ですから、時間をかけて作業を進めることができます。とくに昨今のように電子図書が流通し、活字のデジタル化が進めば、データベース化は容易になるでしょう。

統合版日本民話データベースは、ファジーな構造を持っていますから、音声情報と文字情報、さらには映像や動画の情報を付け加えることができます。問題は、コンピュータのCPUの計算速度とハードディスクの容量ですが、これは、例えば、現在のウインドウズ7のレベルでも十分に可能です。

日本民話データベース作成委員会は、できるかぎり地域的な

偏りを排して、全国にわたる音声情報を入力したいと考えますが、予算や時間の不足によって、今すぐこれが達成できない場合にも、最終的には十分な結果を得られると考えています。

6-2. 作業者による入力作業の偏りを、どうするか

作業者による入力作業の偏りの問題に関しても、あまり大きな問題はないと考えています。それは、やはり日本民話データベースのファジーな構造にあります。たとえば、キーワードや話型の入力ミスは深刻な問題です。しかし、それによってプログラム全体が動作を止めてしまうことはないのです。

こうした誤りは、いつでも修正できますし、不足は追加できます。デジタル・データベースの強みは、常に進化しているということです。最初から「完全なデータベース」を作る必要はないのです。たくさんの人が、一定のルールに従ってデータを自由に使用し、一定のルールに従って、誤りを指摘し、修正してゆけばよいのです。

誤りの訂正は、それほど難しくありません。大切なのは、まずデータベースを正しく設計すること、つぎは使用上のルールを正しく決めること、最後は決められたルールに従ってデータベースを正しく管理することです。

7 民話データベースの国際協力の現状と可能性

つぎに民話データベースの国際協力の可能性についてお話し

します。

二〇一〇年十月、二〇一一年七月の二度にわたり、ソウル近郊の韓国学中央研究院で、口承文芸資料のデータベース化国際協力に関するワークショップが開催されました。昨年十月は日韓、本年七月は日韓中の枠組みで、本年は中国社会科学院のメンバーが新たに加わりました。そこで得た知見を踏まえて、とりあえず先行する日韓国際協力の現状と可能性について報告します。

7-1. 韓国の民話記録

7-1-1. 一九四五年の光復までの民話記録

韓国における民話の記録は一九二〇年代にはじまり、一九二六年に沈宜麟の『朝鮮童話大集』、一九二七年に鄭寅燮の『温突夜話』、一九三〇年に孫晋泰の『朝鮮民譚集』が刊行されます。とくに孫晋泰の『朝鮮民譚集』は二〇年代初頭からの記録で、語り手や記録日時が正確に記された極めて質の高い記録です。その後、秋葉隆らによって、朝鮮半島全域にわたる文化人類学的な調査が開始されると、彼等と行動をともにした任哲宰の調査記録が開始されます。任哲宰の記録は、一九八七年から九三年にかけて『任哲宰全集・韓国口傳説話』全12巻として刊行され、半島全域にわたる貴重な民話記録となりました。（任哲宰は、音声録音の普及した一九六〇年代以降は、韓国本土の民話の調査を行い、その記録はこれから紹介する韓国学中央研究院に寄贈され、データベース化されています。）

7-1-2. 一九四五年の光復以降の民話記録

日本の植民地支配が終わった一九四五年以降は、韓国独自の民俗調査が行われますが、その一方で南北分断という状況下で、私たちの目にするのできる記録の大半は、朝鮮半島（韓半島）南部のみのものとなりました。

一九四五年以降の韓国の民俗記録の大きな特徴は、大規模な調査が国家事業として行われたことです。その第一次調査は、一九七〇年代初頭に文化広報部が韓国文化人類学会に委嘱して行ったもので、韓国の全道を対象に、各年度ごとに一道の調査を集中的に行うという組織的なものでした。この調査の結果は竹田且先生と任東権先生のご努力で翻訳され『韓国の民俗大系』という形で、大部分の成果を日本語で読むことができます。そこに「口碑伝承」という形で納められた記録は大きく民話・説話・謎と諺・方言などに分類され、民話は「説話」として分類され、その扱いは各道ごとにばらつきがありますが、貴重な記録であるといえます。

韓国政府主導で行われた第二次の調査は、一九七八年から一九八四年まで韓国精神文化研究院（現・韓国学中央研究所）が、全国六十の市町村（市・郡・面）で行った『全国口碑文学調査研究』で、その成果は『韓国口碑文学大系』全八二巻に収められています。この調査は、やはり説話と民話という二つのジャンルに分類されていますが、当初から「口碑文学調査方法」という厳密な調査マニュアルが用意され、その成果である一五一〇七話

の説話（＝民話）は、この調査の中心となった趙東一によって話型分類されています。（説話と同時に調査された民話六一八七、巫歌三七六、その他資料二十一も『民話巫歌類型分類集』によって分類整理されています。）

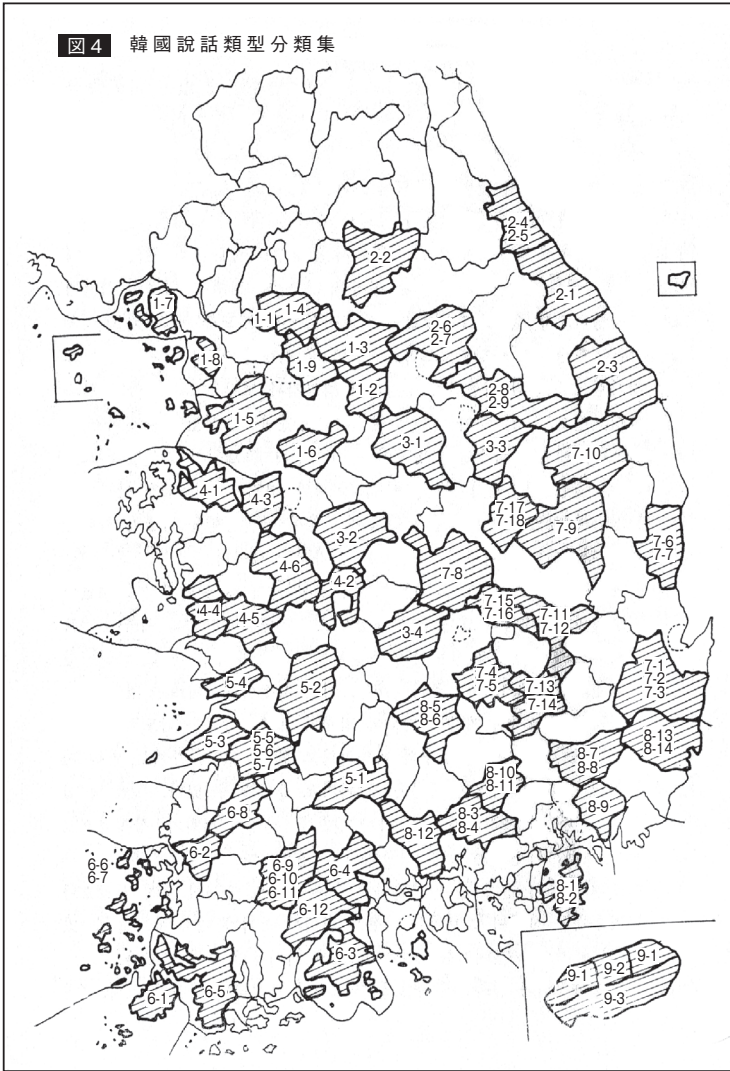
韓国の政府主導で行われた第三次の調査は、二〇〇八年から十年の計画ではじまり、現在も進行中のプロジェクトです。この調査は、前回の韓国精神文化研究院の調査を引き継ぐ形で韓国学中央研究所が主導しているもので、前回調査の欠落部（未調査地域）を補足して、全国すべての地域の記録を完成し、同時にその成果を全てデータベース化するという計画です。（図4の斜線が前回調査地域）

こうした国家的な調査のほかにも、もちろん個人の調査や研究も存在します。

そのもっとも顕著な業績は、崔仁鶴の仕事であろうと思われます。崔仁鶴は、一九六〇年代に日本に留学し、関敬吾の『日本昔話集成・大成』に匹敵するタイプ・インデックスを刊行することによって、韓国昔話の国際比較研究を可能にしました。

崔仁鶴は、自身も韓国各地で聞き取り調査を行い、その成果を日本語でも公表しているので、日本の読者にも親しみ深い存在です。また最近では、李秀子、黄仁徳のように四〇〇話、五〇〇話クラスの話者の記録を刊行する研究者も出現しました。

図4 韓国説話類型分類集



7-2. 韓国のデータベース化の現状

7-2-1. 韓国と日本の民話記録の性格の違い

一九六〇年代以降、音声で記録された民話記録の資料をデータベース化する上で、日本と韓国が世界で有数の可能性を保持

していると考えられます。この二つの国が保有する民話記録の質と量は、現時点では、世界のどこの国にも比較できない水準にあると思われます。

もちろん、文化人類学などのフィールドワークを通じて、自らが調査対象とする人々の語りの記録を量的にも質的にも高い水準で保有している研究

者の数は、少なくないと思われれます。しかし日本や韓国のようなタイプの近代国家のレベルで、これだけの質量を確保することは、今後も難しいのではないのでしょうか。

そして、きわめて興味深いのは、この二つの国の保持するデータの性格が大変異なるということです。すでに述べたことから容易に想像がつくことですが、韓国のデータは、国家的なプロジェクトによって得られたものがほとんどであるのに対

して、日本にはこのような国家プロジェクトが皆無に近く、ほとんど全ての記録が個人に帰属しているということなのです。

こうしたデータの性格の違いは、両国のデータベースの性格の違いに反映されます。

韓国のデータベースは、短期集中のローラー作戦によって生み出されますから、今回の第三次プロジェクトが終了すれば、全国の津々浦々を覆い尽くすことになり、地図上の空白地帯は消滅します。日本の場合には、大東島以外のすべてを調査し尽くした沖縄国際大学の調査をのぞけば、このようなタイプの調査は存在しません。後で述べるように、日本民話データベースの地域的な偏りと空白は、目を覆うばかりです。

しかし、その一方で日本の民話調査は、個人が時間をかけて、同じ調査地を何度も訪れて、丁寧な話を聞きます。その結果、一〇〇話を越す優れた語り手の数は、韓国に比べて圧倒的に多いと思われまゝ。誤解のないように申し上げますが、韓国のデータベースに納められた話を聞いてみると、一人ひとりの語り水準はとても高いし、話を聞く調査者の姿勢も実にしっかりしていると思われまゝ。しかし、ある限られた時間内に、ある地域をくまなく調査する使命を与えられた調査者には、限界があるのではないのでしょうか。

日本の場合には、沖縄国際大学の場合も、じつに丁寧な調査を行い、さらに調査後に地域民話の会を組織して、同一地域の同一の語り手を数次にわたって訪れることによって、多くのすぐ

れた語り手を発見していることを申し添えます。

7-2-2. 韓国におけるデータベース化とその公開状況

つぎに韓国におけるデータベース化の進行状況についてお話しします。

韓国における民話データベース化の作業は、二〇〇〇年の半ばから始められた「韓国歴史情報統合システム(한국역사정보통합시스템)」の一環として着手されました。これは、韓国学中央研究院に付属する「蔵書閣」という図書館の所蔵する古図書、古文書、写真・絵画資料、音声資料などをオンラインで公開する事業で、<http://yoksaakackr/> というURLを入力すれば誰でも資料にアクセスできます。このページの音声資料(음성 자료) という中央のタグをクリックすると、誰でも①韓国口碑文学大系、②韓国方言資料集、③韓国民謡大観という3つの音声資料を利用できるのです。(図5)

この三つの資料は、それぞれ大変優れたものですが、ここでは「韓国口碑文学大系」について紹介します。

このデータベースには、先に紹介した一九七八年から一九八四年までに韓国精神文化研究院(現・韓国学中央研究院)が行った調査データ(説話一、五一〇七、民謡六、一八七、巫歌三七六、その他資料二一)のほか、「韓国口碑文学大系」に収録できなかったデータも公開されています。

これらの語りの資料は、①巻別、②素材(モチーフ)別、③



図5

類型(話型)別、④地域別、⑤地図、⑥題目(話名)別に整理されています。巻別は一目瞭然ですから省略して、素材別から紹介します。

②の素材(モチーフ)は、①天文、②地理、③人間の三分類で、たとえば人間は①通過儀礼、②生き方の二分類で、通過儀礼

は、①誕生、②疾病、③死別、生き方は、①身分、②家族と親族に別れます。この分類は、きわめて簡潔で骨太ですが、たとえば身分に関わる話には五九六件、家族と親族は八〇〇件のほり、あまり実用的ではないかもしれません。しかし、地理に分類されている時にまつわる話は4件のみですから、利用しやすい項目もあるでしょう。

④の地域別、⑤の地図は、たいへん使いやすいと思いますが、地図は現時点では未完成のようです。

⑥の題目(話名)別は、ハングルを入力することで、調査時点での決定題名を入力すれば当該資料を検索できます。

しかし、なんといってもこのデータベースの核心をなすのは、③の類型(話型)別索引です。この独自で複雑な分類に関しては、田畑博士さんのたいへん優れた論文がありますので、その翻訳を利用していただきます。(田畑博士「日韓中昔話比較研究の可能性」in『石井正巳編・韓国と日本をむすぶ昔話』二〇一〇・二／東京学芸大学)

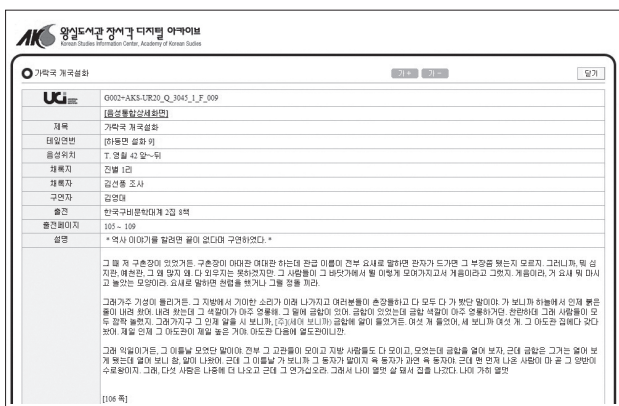
この分類は、すでに述べたように趙東一によって案出されたものです。趙は、韓国の説話(『民話』)を、①勝つことと負けること、②知っているか知らないか、③騙したり騙されたり、④本当と間違ひ、⑤動いたり止まったり、⑥行ったり来たり、⑦良くなったり悪くなったり、⑧続いたり止めたり、という8つのタイプに分けました。そして、たとえば勝つことと負けることを、①勝つべくして勝つ(二二八)、②勝つべきなのに負け



【图 6】

번호	서명	구분자	채택지	유형	서지	음성	텍스트
1	가락국 개국설화 * 역사 이야기를 들으면 재미 없는데 구연하였다. *		김대중 대통령도-영월군-영동면	설화	서지	음성	텍스트
2	장흥의 장해막 * 장흥 장수 이야기를 계속 듣자, 이 지방은 아니고 장흥 이야기라면서 김진중장흥도-영월군-영동면 설화		김진중 장흥도-영월군-영동면	설화	서지	음성	텍스트
3	거인 암사 폭이 그 폭과 제자되어 폭속 구한 이 순신 * 이순신 장군이라든가 임유신 장군등 장군 이야기를 풍자하 다음에 잘 없기북강원도-영월군-영동면		북강원도-영월군-영동면	설화	서지	음성	텍스트
4	가죽도에 초렁이가 있는 이유 * 잔손이 풀이내가 오랑이풀이 같은 새나이를 이야기하자 폭이 이 이야기가 생각난 모양이다. 알 이야기가 끝나자 아내 이 이야기를 시 길쭉약 강강설화-거제군-동부면		강강설화-거제군-동부면	설화	서지	음성	텍스트

【图 7】



【图 8】

491년을 갖다가 정다산론에 보면 591년이라 그랬어. 여져 머른 말씀도 해 놓고 했으니까, 머느 말씀이 맞는지 남이 욱한 것도 써 놓고 남이 업신 여긴 것도 다 기록하

듣기

Q_3045 41:32

【图 9】

る(八四)、③負けるべきなのに勝つ(二四四)、④負けるべくして負ける(一二七)、の四つに下位分類します。この1のタイプは、田畑さんによれば「神霊や英雄が勝負する話。神人、建国始祖、英雄、将師などが主人公」となる話で、合理的といえは合理的な分類です。(图6)

たとえば、この分類の①(勝つことと負けること)の(勝つべくして勝つ)の冒頭に位置する「駕洛国(金官伽耶) 開国説話

가락국 개국설화」を見てみましょう。(图7)

ここには、いくつかのタグがありますが、一番左の(テキスト)というタグをクリックすると、图8のようなページが表示されます。そこには、話者、記録者、記録地、テープ番号、記録日時などの情報が記されています。(图8)

そしてなにより便利なのは、「口碑文学大系」に収録されたテキストが音声とともに全文収められていることです。

ページを下までスクロールすると、左のような音声用のアイコンがあり、これをスタートさせると「襍洛国（金官伽耶）開国説話」を語る金ヨンデさんの声が流れてきます。（図9）

8 韓国・東アジアとの共同作業の課題

ATというスタンダードの尊重

ここまで、日本と韓国の民話調査とデータベース化の現状について紹介してきましたが、ここで「東アジア、とくに日本と韓国、そして中国の間で、民話の資料を共有し、データベース化を推進するにはどうしたらよいか」という見直しについて、私の考えを述べさせていただきたいと思えます。それは、難しいことではありません。

8-1. 国際比較の視点の必要

韓国であれ、日本であれ、自国の民話のデータベースを作成しようとする場合に大切なのは、まず自分たちの国の文化遺産である「民話」を記録し、保存し、研究や教育に役立てようということでしょう。とくに一九六〇年代以降に記録されたアナログ・データは、時間とともに消滅する運命にあります。この時代の記録をデジタル化することによって半永久的に保存し、データベース化することによって実用に耐えるものにするには、現代の民俗学研究者の責務です。

しかし、誰でもすぐに気がつくことですが、民話は世界的

な広がりを持っており、簡単に国境を越えてしまいます。たとえば、「シンデレラ」はヨーロッパやアメリカで人気の高い話で、子どもたちなら誰でもディズニーのアニメは知っているし、ディズニー・ランドのシンデレラ城が大好きです。しかしこの話の一番古い記録は、周知の通り、中国の唐時代の『西陽雜俎』にあります。また、韓国には「コンジ・パッジ」(KTCJ0)があり、日本には「米福・粟福」(ST205)等があります。

民話には、同じような問題が山のようにあります。韓国民話の特性について理解しようとする時には、中国や日本だけでなく、ヨーロッパやアメリカやアフリカや中近東やロシアやオーストラリアのことを考えないわけにはいけません。自分たちの国の伝統文化の独自性を理解しようとするれば、東アジアだけでなく全世界に広がる民話・民俗のネットワークと連携する国際比較の視点が必要です。

8-2. 分類の問題

日本に柳田國男、関敬吾、小沢俊夫、稲田浩二などの、それぞれ特徴ある民話の分類案があるように、韓国には崔仁鶴、曹喜雄、趙東一などの優れた分類案があります。中国にも、丁乃通、金榮華などの分類案があります。

私は、民話データベースを作成するにあたっては、それぞれの国の、それぞれの事情に従って、複数の分類案を並行して採用すればいいと考えています。しかし、どうしても忘れて欲し

くないのは、国際比較の視点です。さまざまに批判のあることは承知の上ですが、すでに国際比較の上で「de facto standard」として確立しているアアルネ・トンプソンの分類（A・T・A・T・U）を尊重し、必ず検索項目に付け加えることが大切です。

そして、ことに韓国の場合には、国際比較とA・T・A・T・U分類に当初から配慮した崔仁鶴の話型分類（K・T）が存在しますから、この分類だけは必ず検索項目に付け加えて欲しいと考えています。

韓国の人たちが、韓国の民話を韓国固有の仕方でも分類するのは、当然のことです。韓国の民話データベースは、まず韓国人の研究者や教育の現場で生かされなければいけませんから、そのために韓国人や韓国文化の専門家だけが利用できる検索項目を、いかに詳細に構築されても構いません。しかし、その一方で、韓国語が十分に理解できなくても、韓国の文化や伝統に関心のある私のような外国人の研究者や一般人が、誰でも利用できることが大切だと思います。

たとえば日本の場合には、すでに多くの韓国民話に関する資料を日本語で利用できる体制が整っていますから、すでに多くの人が韓国民話に親しみを持っています。英語圏にも韓国民話の資料は少なくありません。ですから、ぜひ、民話の国際比較という視点を配慮していただければと考えます。

将来、インターネットを通じて、韓国の民話を世界に発信される場合にも、必ずA・T（A・T・U）の参照が必要になるでしょう。

8-3. 検索項目の共通化

日本民話データベースと沖繩伝承話データベースの検索項目は、現在三十三あります。しかし、このシステムをさらに進化させる為に、たとえば文字情報とか動画情報とか、さまざまな検索項目を付け加えることは難しくありません。

それは、このデータベースの基本構造が、エクセルによって組み立てられているからです。現在、このデータベースにおいて最も大切なのは音声情報ですが、これに文字情報や動画情報を追加したり、新しい民話の分類案を追加したりすることは、技術的には難しくありません。データベースの構造は、非常にシンプルなのです。

ここで、もう一度、三十三の項目を振り返ってみましょう。

この三十三の項目のうち、「③ 決定題名、④ 話者がつけた題名、⑤ 話者名、⑦ 生年月日、⑧ 性別、⑨ 調査時話者住所、⑩（話者の）出身地、⑪ 記録日、⑫ 記録場所、⑬ 採訪者名、⑭ 記録者の所属組織、⑮ 17分類、⑯ 大成名（K・T話型名）、⑰ 大成番号（K・T番号）、⑱ A・T番号、⑳ 発句、㉑ 結句、㉒ あいづち、㉓ 伝承事情、㉔ 文字化資料、㉕ 備考、㉖ 写真、㉗ キーワード、㉘ 梗概」などは、番号はともかく、韓国でも日本でも大切な情報であろうと思われまます。「⑰ 分類」という項目については、これまで言及しませんでした。⑰「分類」という項目については、これまで言及しませんが、日本の記録テープには、昔話のほかに伝説、神話、世間話、民話、俗信、生活上の経験譚など、さまざまの

ジャンルの話が登場します。それを全て切り捨てずにデータ化しますから、その分類が必要になるわけです。韓国の場合にも、おそらく調査テープには、さまざまなジャンルの口頭伝承が含まれると思いますから、とりあえずは「⑩分類」のような形で整理しておくと思います。

幸い、韓国・中国・日本は漢字文化圏です。したがって、検索項目には、まずそれぞれの国の漢字を利用すればいいと思います。さらに、将来の世界的な展開を準備すれば、英語の検索項目の共同利用も視野に入れておきたいと考えます。

もし、今後のデータベース作成上で韓中日三国の共同作業が始まれば、互いに、これらの項目の一つ一つに関して検討会を設ければいいと考えます。

今後のデータベース作成上で韓国と日本、そして中国の共同作業が始まれば、双方で、これらの項目の一つ一つに関して検討会を設ければいいと考えます。

まとめ

民話データベースの作成とその国際化にあたって発生する問題を具体的にお話したために、たいへん長い話になって恐縮です。

しかし、私の話の内容は極めて短く、次の八つに要約出来ます。

- ①音声データの保存には、Wave形式を用いること。
- ②音声データの一件ごとの切り分けには、熟練が要るので、チームを作り、責任者を決めて、慎重に行うこと。
- ③作成チームの責任者は、メンバーの作成したデータを十分にチェックし、メンバーに対する指示を徹底すること。
- ④データの整理には、エクセルを用い、共通のフォームに従って記入すること。
- ⑤データの記入項目は、最初に時間をかけて、十分に準備すること。(後で項目を追加することもできるが、手間がかかるので、最初にしっかり決める方がよい)
- ⑥データ検索の項目は、国際化に配慮して、まず漢字を使用し、将来的には英語の併用も考慮すること。
- ⑦話型分類に、国際比較の視点を必ず取り入れること。
- ⑧以上のように、周到な努力を重ねても、「完全無欠なデータベース」は絶対にできない。しかし、データの追加や修正は、いつでも可能だから、失敗を恐れずに、まず作成に着手すること。報告は以上ですが、二〇二二年四月以降には約五万件のデータ検索が可能な統合版日本民話データベースが完成する予定です。この検索CD-ROMを、研究・教育目的の使用にかぎって無料配布いたします。ご関心のおありの方は、樋口淳まで御連絡ください。アドレスは higuchi007@hotmail.com です。

(ひぐち・あつし／専修大学文学部)