

公益財団法人国土地理協会 第 15 回学術研究助成

ブータンのゾンカ語表記地図製作に関する調査研究

研究成果報告書

研究代表者：平山 雄大（早稲田大学平山郁夫記念ボランティアセンター講師）

共同研究者：高橋 洋（日本ブータン研究所研究員）

共同研究者：西田 文信（早稲田大学教育・総合科学学術院准教授）

目 次

1. 背景	3
2. 研究目的	4
3. 研究対象（ブータンにおける地図出版の基本状況）	5
4. ゾンカ語環境の構築方法	8
(1) 作業環境	8
(2) ゾンカ語文字コード	8
(3) ゾンカ語組版	9
(4) ゾンカ語フォント	10
5. 地図作成の作業手順	11
(1) ゾンカ語正書法、ローマ字化手法の確認	11
(2) ゾンカ語地名集（Gazetter）の構築	12
(3) ゾンカ語地名の入力とマッピング	14
(4) 地図出力	14
6. 研究成果	15
(1) 多言語地名テキストデータベースの構築	15
(2) 成果物としてのゾンカ語表記地図の試作	15
(3) 地図製作技術のローカライズに関する諸問題の検討	15
(4) 社会言語学	17
(5) 言語政策・文化政策の基礎資料の収集	18
7. 考察	18
8. まとめと課題	19
ゾンカ語綴りの確認のために参照した主な資料	22

1. 背景

信頼できる地図の存在は、学術的な地域研究はもちろんのこと、開発や自然保護といった分野においても作業の前提条件となる。しかしながら、隣接する中国・インドとの国境未画定地域に囲まれたブータン王国（Kingdom of Bhutan、以下ブータン）の場合、そのような用途に耐え得る信頼できる地図はごく一部地域を除いては刊行されていない。また、ゾンカ語（Dzongkha）という国語（公用語）を持ち伝統文化の保護を国是としながら、主に技術的な理由により既存の地図はすべて英語表記である。

ブータンは言語的に多様性に富んでおり、確認できる限り 18 言語が分布している（図 1 参照）。これらの諸言語は公的には文化遺産の保護を謳って保存されるべきと考えられているが、実際はほとんど研究されていないといっても過言ではない。近年ブータンでは国語（公用語）たるゾンカ語の普及が図られて来ており、ゾンカ語以外の諸言語を取り巻く状況は急速に変貌を遂げつつある。

ゾンカ語は、17 世紀以来のンガロン（sNga-slong）方言を母体として形成されたもので、標準語として政府が使用を奨励している言語である。Drukke、Drukha、Dukpa、Bhutanese、Jonkha、Lhoke、Lhoskad、Hloka、Lhoka、Bhotia of Bhutan、Bhotia of Dukpa、Zongkhar、Zonkar と称されるこの言語は、語源から言うと「ゾン」（城塞僧院）で使用される「カ」（口＝言語）を意味し、それを母語とするのはブータンの総人口の 20～30%とされている。ブータンの学校教育はゾンカ語を除くほぼすべての科目が英語で行われているため、小学生でも流暢な英語を話すが、英語が普及しすぎることによって自

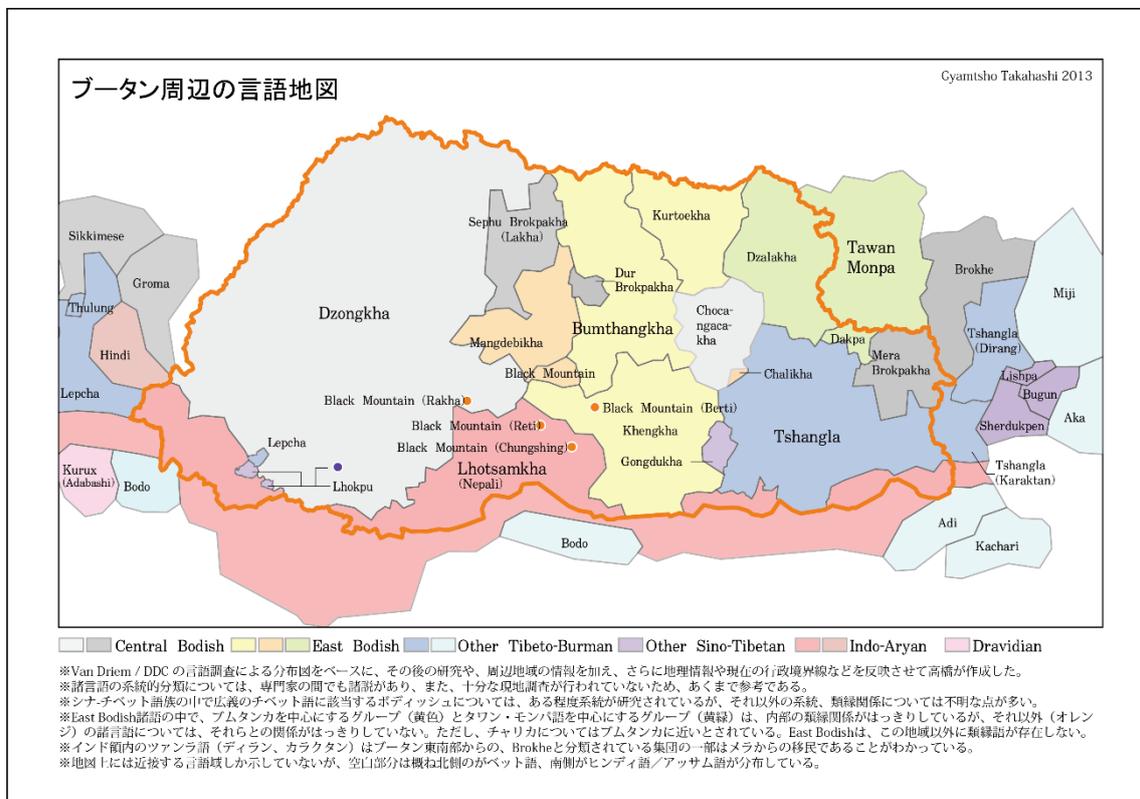


図 1 ブータンの言語分布

国の文化が廃れることを危惧する政府によって、近年は積極的なゾンカ語の普及政策が行われている。

国家統合を目的としてゾンカ語の普及を促進するため、ゾンカ語開発委員会 (Dzongkha Development Commission: DDC、以下 DDC) がゾンカ語の標準化、識字の奨励、辞書編纂、教科書編集等に従事している。しかしながら、主にマンパワー不足及び資金面の問題で満足いく成果がまだまだ得られていないのが現状である。目下ゾンカ語の綴りの統一化が図られているが、全国的に普及するにはなお時間を要する。ゾンカ語の問題点としては以下の点が考えられる。

- [1] ゾンカ語の定義は曖昧であり、「フォーマルな場で用いられる文章語としてのゾンカ語」、「洗練された話し言葉としてのゾンカ語」、「ブータン西部のローカル言語としてのゾンカ語」という少なくとも3つの意味を有する。
- [2] ゾンカ語はマスコミ (ラジオ・テレビ・新聞等) の言語、伝統歌謡・伝統舞踊の言語と認識されているが、法的文書や契約文書はより古い言語的特徴を有する古典チベット語 (Chöke) で書かれるのが一般的である。
- [3] ゾンカ語は公用語ではあるが、母語話者数は総人口の20~30%に過ぎず、国内ではツァンラ語 (Tshanglakha)、ネパール語 (Lhotsamkha)、英語を含めた4つの lingua franca が話されている。
- [4] 古典チベット語は5母音体系であるが、ゾンカ語は8母音体系の言語である。保守的な綴りと革新的な発音との間に乖離が起こっており、表音的表記が難しい。
- [5] ブータンでは英語がかなりの程度操れないと就職が困難である。ゾンカ語能力を活用することのできる就職先は多くなく、ゾンカ語教師などに限定される。
- [6] 多言語・多民族国家を国民国家へと統合させるゾンカ語教育政策は、上手く機能しているとは言い難い。

このような状況に加え、ブータンにおいて地図は国民にとって身近な存在とはいえ、そのことが自国語地図製作へのモチベーションの欠如に繋がるという悪循環が展開されている。そもそも現在に至るまで出版文化・出版産業の蓄積が乏しい人口約75万人の同国が、自国語による本格的な地図を独自に製作することは経済的、技術的、人材的にも困難であった。

2. 研究目的

本研究はブータンの国語 (公用語) であるゾンカ語による全国地図の実製作を行い、その有用性及び技術的な課題を検討・評価することを目的としている。

出版文化や技術の蓄積が乏しいブータンでは、これまで本格的な地図はすべて英語を記述言語として作成されてきた。これには、英語版地図の社会的ニーズが大きいという理由もあるが、複雑な文字処理が必要になるゾンカ文字による高度にビジュアル化された出版物を製作するための技術的蓄積が存在せず、また地図製作技術そのものも諸外国の技術指導のもとでなされてきたため、技術移転が容易な英語が選ばれてきたという理由も大きい。

一方で伝統文化の保護と国土やその自然環境の保全是、ブータンのナショナル・アイデンティティの、そしてそれを重視する国民総幸福（Gross National Happiness: GNH）政策の基幹をなすものであり、国語による国土地図の不在はそれに逆行するものである。本研究ではオープンソースの地理空間情報システム（GIS）やリモートセンシング、多言語テキスト処理といった最新の IT 技術を応用して、本格的なゾンカ語地図を製作するための基礎技術を開発し、その可能性と課題を実証することを目指す。

これまで、ゾンカ語、ゾンカ文字を使用した地図が事実上存在しなかったため、それが英語地図に比べてどのような意味をもち、また実際にどのような社会的ニーズがあるのかといった基本的な位置づけについても未知の状態にあった。近代国民国家の共通語としてのゾンカ語の歴史は始まったばかりであり、言文一致、正書法の確立といった基本的な側面においても未解決で開発過程の部分が大きい。地図の実制作という作業を通して、ゾンカ語の現状及びブータンの言語事情やその課題を明らかにすることも本研究のねらいのひとつである。

また、ゾンカ語は国民国家の統合の象徴としてブータン人全体に支持されているが、ゾンカ語を母語とするブータン人は全人口の 20~30%ほどであり、ブータンは基本的に多言語社会だといえる。ゾンカ語を母語としない地域では、ゾンカ語地名、現地語地名が別に存在し使い分けられていることも多い。南部のインド国境付近の町では歴史的・慣用的な呼び名があり、古典チベット語で記述された中世文献資料の地名表記は現代ゾンカ語表記とはまた異なる。コンテキストによって同じ場所の地名が変わるため、先行研究でも資料に示される場所の比定が難しく、またその結果さまざまな混乱や誤解が生じることが多かった。GIS 上に地名データベースを構築することにより、複数の言語、複数の表記方法による地名を統合的に扱うことで、この問題の解決の糸口を得ることも本研究の視野に入っている。

3. 研究対象(ブータンにおける地図出版の基本状況)

ブータンでは、国語であるゾンカ語を表記言語とした詳細な全国地図は現在に至るまで公刊されたことがない。正確に言えば、本格的なゾンカ語地図自体が存在しない。現在一般に入手可能な全土をカバーする詳細地図としては、農業省（Ministry of Agriculture: MoA）による *Atlas of Bhutan 1:250,000: Land Cover & Area Statistics of 20 Dzongkhags*（1:250,000/1997 年）や *Farm Road Atlas of Bhutan 2012*（1:80,000/2012 年、**図 2** 参照）、土地管理委員会（National Land Commission: NLC、以下 NLC）による *Atlas of Land Cover and Institutional Facilities*（1:100,000/2016 年、**図 3** 参照）などがあるが、これらはすべて英語表記（タイトル、解説、凡例なども含む）である。このことは背景で述べたようなブータンの特殊な言語事情から、英語地図の社会的ニーズがゾンカ語地図よりも高いことが最大の理由だと考えられる。

前例がないため、英語表記、ゾンカ語表記による地図の利便性や問題点はこれまで比較検討されたことがなかった。しかし、アルファベットを使った転写（transcription）、つまりゾンカ語のローマ字表記では、元のゾンカ語表記の情報の一部が失われ、その結果語義が分からなくなったり、異なる地名が区別できなくなったりするといった問題があること

が指摘されている。漢字地名をローマ字表記や仮名表記にすると同音異字の地名が判別できなくなったり、英語地名を仮名表記にすると L 音と R 音の区別がつかなくなったりするのと原理的には同じ現象である。またもうひとつの潜在的な問題は、ローマ字化した地名が英語読みされてしまうという可能性である。

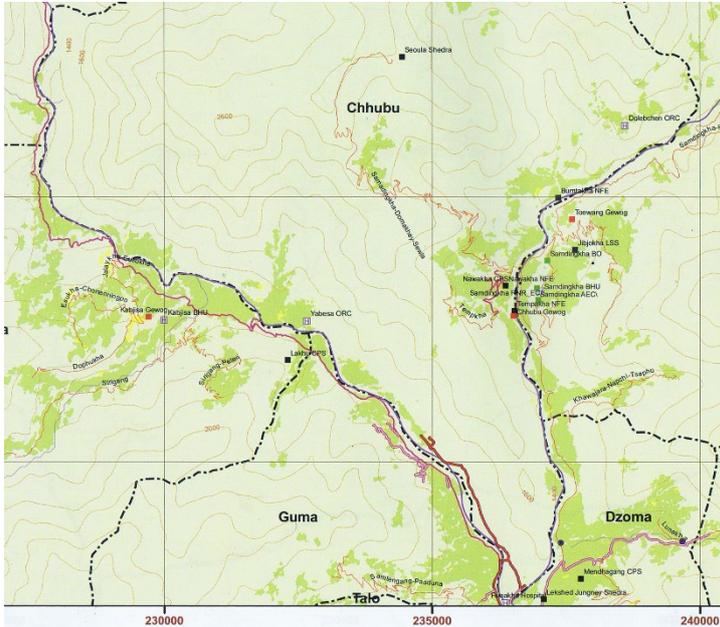


図 2 先行・比較事例

Farm Road Atlas of Bhutan 2012 (農業省、2012 年)

道路局 (Department of Roads: DoR) が管理する舗装された幹線道路以外の地方道は農林省 (旧・農業省) によって管理されている。この地図帳はいわばその台帳である。基本は 1:80,000 だが、地図範囲が緯度・経度ではなく 1 郡または近隣の複数の郡なので、郡の面積によって縮尺が異なる。

地名は郡名のみが掲載され、集落名などはない。そのため、道路地図であるものの分岐点や終点の地名が特定できない。施設名としては病院や学校などの政府機関の位置だけが示されている。英語表記についてはブータン政府の一般的な綴りを採用している。

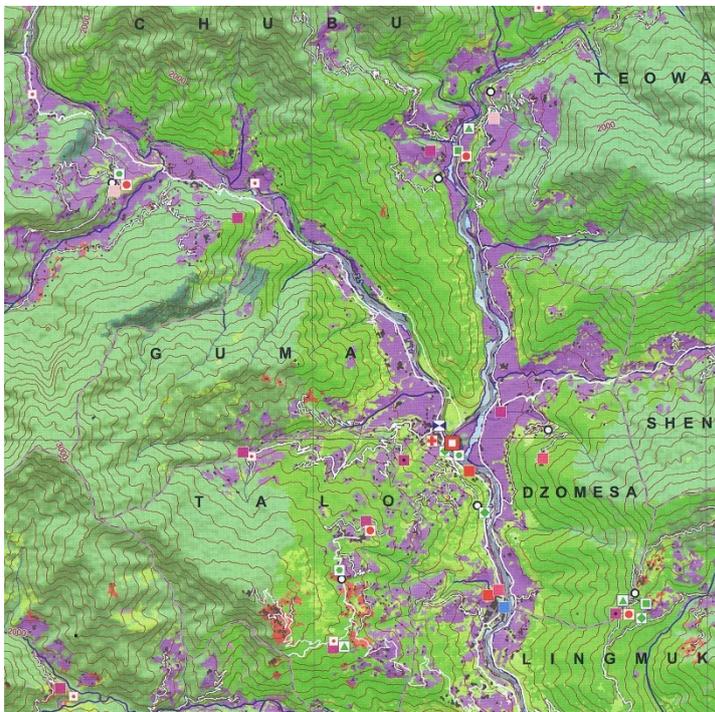


図 3 先行・比較事例

Atlas of Land Cover and Institutional Facilities (NLC、2016 年)

全国土地利用調査 (Land Use and Land Cover of Bhutan: LULC 2016) をベースにした地図帳で、ブータン全土を同一縮尺 (1:100,000) でカバーした地図帳としては初のものである。地名は基本的には郡名までだが、一部の地域では集落名までカバーしている。またこれも一部の地域ではあるが、山名・河川名といった自然地形名が比較的豊富に掲載されている点は、いままでの同種の地図にない特徴である。

施設は病院や学校など政府機関のみ記号で示されている。施設名は記載されていない。この点、地名 (集落名など) を掲載しない代わりにそれらの施設をランドマークとして使う農業省の地図とはまったく逆の発想にある。地名表記は一般的な政府文書とはかなり異なった部分がある。

発音の再現性以外にもいくつかの問題がある。ゾンカ語では音節の区切りと単語の区切りが明示されるのに対して、アルファベット表記では音節の切れ目が表現できない。例えば、モンガル県は英語表記では「g」の連続を嫌って“Mongar”と表記されるのが一般的だが、ゾンカ語では“mong sgar”と表記される。しかしこの英語表記方法では“mong sgar / mon sgar / mo ngar”のようなゾンカ語が同じ“Mongar”になってしまうことになり、区別できない。地名としての発音が同じであっても、語源が識別できるように表記にバリエーションをもたせた、いわば「同音異字地名」が英語（ローマ字）表記では区別できなくなるのである。この点はチベット語翻字（transliteration）として一般化しているワイリー方式と大きく異なっている。

ゾンカ語地名表記の利点が明白でありながら普及しなかった理由は、社会ニーズだけではない。ブータンの地図製作の歴史は独立までインドの宗主国であった英国の植民地行政や、その後の中印紛争といった歴史的な政治的・外交的状況と密接に結び付いていた。一方、チベット文化圏には地図の作成・利用の伝統がなかった。チベットとインド（イギリス）の国境画定を図ったシムラ会議でも、資料として用いられたのは英語地図の地名に手書きでチベット文字を書き加えたものであった（図4参照）。近代化以降も、チベットの地図は中国語（漢字）で、ブータンの地図はインドや欧州諸国で専門教育を受けたブータン人、言い換えればゾンカ語よりも英語が得意な技術者が英語で製作することになり、チベット文字（ゾンカ文字）による本格的な地図は編纂されなかった。

歴史的経緯はともかく、現状としてブータン人自身にとっても地図といえば英語表記が常識であり、極端に言えば「ゾンカ語地図」は見たことがないだけでなく、想像外の存在だというのが実状である。

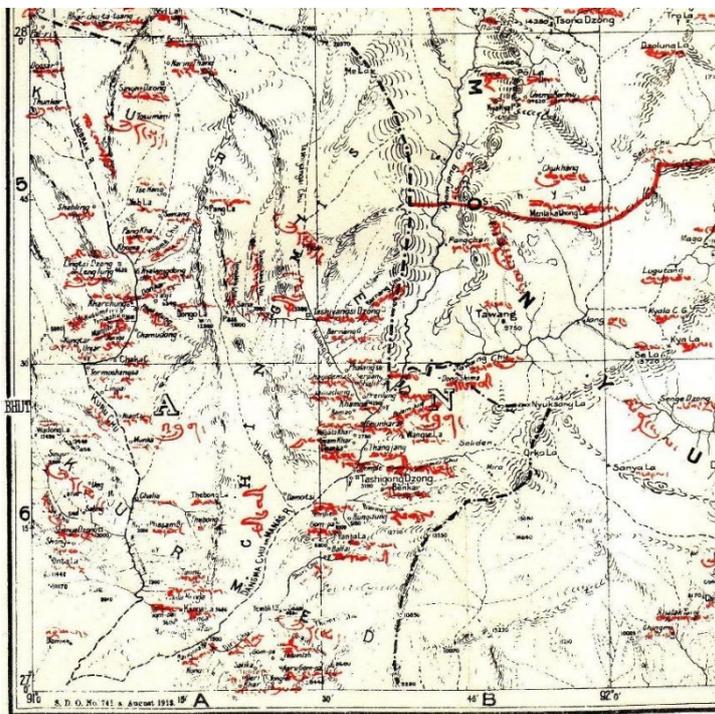


図4 シムラ会議で用いられた地図

注)赤字が、手書きで書き加えられたチベット文字である。

4. ゾンカ語環境の構築方法

(1) 作業環境

地図製作の基本的な手法としては、オープンソースの地理情報システム (QGIS) を利用した既存の地図製作環境 “Gyamtsho Atlas Bhutan” のゾンカ語ローカライズという方法を採用した。なお、同システムは共同研究者である高橋が日本語・英語表記ブータン地図を作成するために開発したものである。

QGIS 上では地名はデータベース管理されるため、ローカライズは基本的にゾンカ語 (ゾンカ文字) 表記地名フィールドの追加と、出力時のデザイン表現の設定だけで可能である。ただし、すべての英語地名に対応するゾンカ語表記データを一括で入手することは難しく、また前例がないため、実際にゾンカ語化されたデータを使ったテスト出力を繰り返して最適化を図る必要がある。そのため、出力の際にある地点・領域に対してゾンカ語表記地名データがあればそちらを、なければ英語表記地名を使用するといったスクリプトを開発して、段階的なゾンカ語化を行った。

なお、英語表記の地名からゾンカ語表記を機械的に復元 (ゾンカ文字列に置換) することは原理的に不可能であるため、ゾンカ語表記データは個別に入手する必要がある。一方でゾンカ語表記のローマ字化は、一定のルールをスクリプトとして記述することによって、ある程度テキスト処理で得ることができる。ゾンカ語表記からのローマ字化スクリプトを開発し、ゾンカ文字表記地名の英語表記地名との照合作業や校正の手段とした。

なおゾンカ語化を段階的に進めていく段階で、既存の英語資料だけでは不可能だった地名の比定が可能になったり、参照した地図や英語版製作の作業過程でのさまざまなミスが判明し解決されたりしたため、ゾンカ語化作業は結果的に英語版地図の精度の向上にも貢献することになった。

(2) ゾンカ語文字コード

ゾンカ語のデジタル化は 1980 年代より進められている。ゾンカ語の表記はチベット文字による現代標準ゾンカ語表現の音写である。この点では古典チベット語や現代チベット語と異なるが共通点も多く、基本的には同じ手法でテキスト処理を行うことができる。ただし、一部チベット語には存在しない複合文字や記号が存在するため、チベット語テキスト処理システムの流用では完全にカバーできない。また、キーボードのマッピングもチベット語とはまったく異なる。初期のゾンカ語文書処理環境は多言語環境の整備で先行していた MacOS 上にチベット語テキスト処理システムをローカライズするかたちで構築され、新聞発行などに用いられた。Windows10 以降は OS にロケールとしてゾンカ語 (dz-BT ロケール) が追加されたため、文字コードは Unicode (UTF-8) ベースになった。それ以前のゾンカ語環境は環境依存が大きく、コード体系が異なるためテキストの再入力が必要な場合が多い。ブータン政府内でも導入されているシステムによって部署ごとに環境が異なるため、ゾンカ語文書の交換の際には Word 文書や PDF のような一般的なデータ形式であっても、トラブルが発生することが日常的であると聞いている。

Web 上のデータでは文字コード体系による文字化けは少ない。ただし、公開されている情報は英語がほとんどであり、例えば各県の公式サイトで言語をゾンカ語に切り替えても

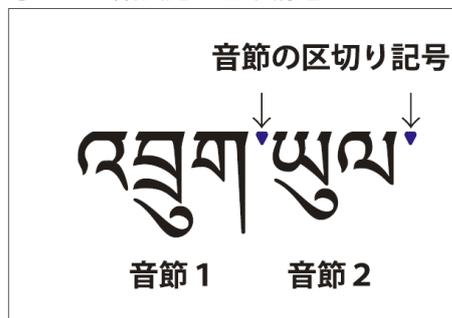
タイトルとメニューの表示文字が変わるだけで本文は英語のままという場合が多い。これは中央政府の省庁のサイトについてもほぼ同様である。

(3) ゾンカ語組版

ゾンカはブラーフミー系の文字なので結合文字を多用し、複雑な文字配置が必要になる。

図5は「ブータン」のゾンカ語名である「ドゥク・ユル」をゾンカ語表記した場合の例である。ゾンカ語（チベット語）表記では単語間の区切り記号はないが、音節間の区切り記号（ツェグ）が明示される。各音節はその音節の基本となる子音（基字）と母音の組合せの前後に前置字・後置字が配置される。

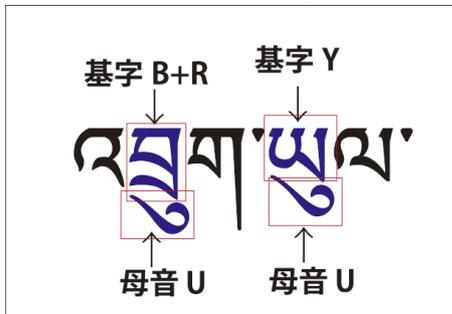
① ゾンカ語表記の基本構造



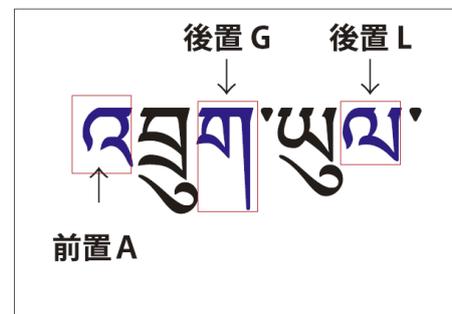
② 語義的な構造



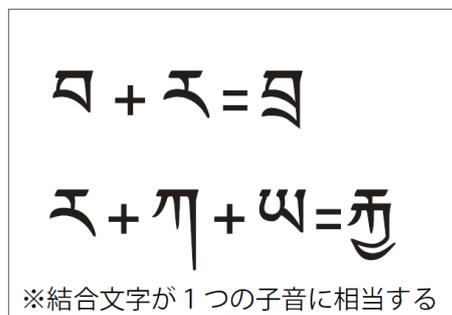
③ 母音と結び付く子音を示す基字



④ 前後の文字は発音の変化を表現する



⑤ 結合文字による字形変化



⑥ 文字の配列方法の概念図

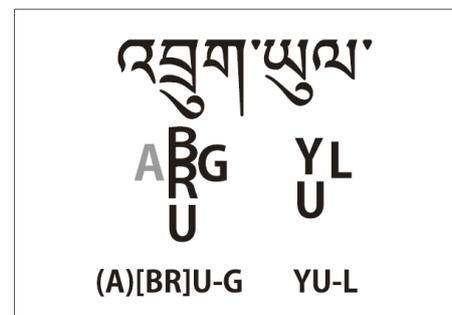


図5 ゾンカ語表記の構造

文字の配置は基本的には左から右へと流れるが、母音は対応する子音の上下に配置され、また2つ以上の子音文字を組み合わせる子音の場合は上下に配置するため、英語や中国語のような文字順が自明な（リニアな）言語と違い、音節の構造によって文字の配置される位置や順序、場合によっては字形が変わる。また、この例の「𑄎」は概念的には「𑄏」と「𑄐」が上下に重なった字形だが、この場合「𑄐」は単独で出現する場合とは大きく字形が変わり「𑄑」となる。また「𑄎」は日本語の続き文字のような一体化した字形で表現される。

このため組版には結合文字や子音・母音の組合せを一組として文字コードを当てる方法と、文字入力の際に位置指定を行う、あるいは入力された文字の連続からコンテキストに応じた配置を行う方法が考えられる。Windows10のロケールでは後者の方法が採用されている。ただしこの場合、連続して配置される文字のすべての組合せをロケールがルールとして管理する必要がある。つまり、例外的な配置（字形）が出現した場合、外字のような形で処理できず、前後の字形まで崩れてしまう。そして地名や人名のような固有名詞の場合には、そのようなイレギュラーな字形がしばしば出現する。一例を挙げると、サンسكريット由来の仏教用語や人名、外来語由来の地名などでは、それを表現するために通常使用される字形とは異なる字形が用いられる。このような特殊な字形はまだUnicodeによって完全にカバーされていないとされる。製作作業でも一部そういった特殊文字、異体字と思われる例が発見されている。一方で単純な誤植もかなりの頻度で存在しており、特殊な例なのかミスなのか判別しがたい。また、こういった場合の包摂のルールも標準化されていないため、現状では対処が困難である。

デザインで問題になったのは、単語の切れ目（英語のスペース）が自明でない点である。ゾンカ語表記では「××（地名）技術訓練校」といった複合語もすべて切れ目なしの一語として扱い、また「技術」のような単語、特に外国語の意識である新語はそれ自体が複合語として表現されることが多い。その結果語長が長くなりがちで、地図上の配置の制約になる。かといってどこで改行を入れるかは語意を把握したうえでないと、つまりゾンカ語リテラシーがないと判断できない。

(4) ゾンカ語フォント

ゾンカ語のフォントとしてはDDCが“DDC Uchen”及び“DDC Wangdi”を公開しており、本研究でもこれを中心に利用している。Uchenは名前の通りチベット語の一般的な書体であるウチェン体で、デザインの的にもチベット語フォントと変わりがない。Wangdiはブータン独自の書体とされるlho yig（ロ・イク）、直訳すれば「南方書体」で、書体としてデザイン的な違いがあるだけで文字としてはウチェン体と1対1で対応し、結合文字など配置に関する規則も変わらない。

ブータンでも一般的な出版物や標識などではウチェン体が用いられることが多い。一方、手書きのノートなどにはロ・イク体が用いられる。この点、日本語の明朝体と楷書体と似た部分がある。地図書体としてはウチェン体の可読性が高い。また、ウチェン体についてはDDCの提供しているフォント以外の選択肢（チベット語フォント）もあるため、地図表現上はウチェン体が扱いやすい。しかしブータンの文化的伝統に沿った地図製作技術を模索するのが本研究のテーマであるため、あえて行政区画名、集落名などの地名はロ・イ

ク体、施設名などはウチェン体という使い分けを試みた。ゾンカ語文書では強調やデザイン的な意図で書体を使い分けるといった習慣がないので、これ自体が実験的な試みである。

通常のゾンカ語文書では数字もチベット文字を用いる。“DDC Uchen”及び“DDC Wangdi”では数字の文字コードにはアルファベット数字の字形が割り当てられているため、標高など数値の出力が必要な場合にはその部分だけ“DDC Uchen Numbers”を用いた。この書体は、数字にアルファベットではなくゾンカ文字が字形として割り当てられた数字書体である。

5. 地図作成の作業手順

(1) ゾンカ語正書法、ローマ字化手法の確認

口語としてのゾンカ語の成立は、遅くとも17世紀以前に遡ることができる。一方、1986年にDDCが設立され現代ゾンカ語正書法の開発を開始する以前は、文語として古典チベット語が利用されていた。歴史が浅く慣用と呼べるほどの蓄積がないこともあって、ゾンカ語表記の正書法は確立しているとはいえない。また、歴史的にゾンカ語が母語であった地域は国土のごく一部にすぎない。チベット語及びゾンカ語を語源とする地名が大多数であるものの、別の地域言語のゾンカ語転写、あるいは音の似たゾンカ語に置き換えられたと思われる地名が多数ある。19世紀後半以降、インド、シッキム、ネパールから多数の民族が移民として定着した南部諸県では、ほとんどの地名がそれら移民の母語によるものであり、ゾンカ語の地名表記はそのゾンカ文字による音写に過ぎないことが多い。このようなケースではゾンカ語表記の地名はゾンカ語話者にとっても、元の言語の話者にとっても語源や語義が了解不能なものになりがちで、当然綴りにも大きな揺れがある。

本研究はゾンカ語表記による地図製作を最終目標としているが、ゾンカ語表記の先行する地図が存在しないため、ゾンカ語地名を地図上に配置するためには英語地図を参照し、英語地名をゾンカ語地名に置き換えるという手法を採らざるを得ない。しかし、ゾンカ語の英語（ローマ字）表記正書法（Romanization）についてはさらに問題が多く、ゾンカ語地名と英語地名の対応に困難が生じる。

このような状況は地図製作に限らず、日常生活でも不都合が生じるのはブータン人も認識しているが、批判はあっても改善はされず、省庁ごと、また筆者ごとに異なる地名表記基準を採用するという状態が長く続いてきた。例えば2016年6月の日刊紙クエンセル（Kuensel）の記事では、前述の農業省とNLCの地図帳では全国20の県（Dzongkhag）のうち3、その下のレベルにあたる205の郡（Gewog）のうち101の綴りが違うと批判している。しかし、そのクエンセルも記者によって地名表記が異なっているのが実状である。ゾンカ語表記の揺れは英語表記ほど大きくはないが、特に非ゾンカ語地域では音写による揺れが大きくなる。現地発音を忠実に表記するか、ゾンカ語（チベット語）の語源に沿った表記とするかという2つの方向性を収束させることができないのがゾンカ語表記の揺れの大きな原因のひとつである。ゾンカ語の地名正書法、ローマ字化手法はいくつかの方法が提唱され混用されている状況で、いまだ標準化されておらず、その中でひとつを選んで表記統一することも事実上不可能という状況にある。

(2) ゾンカ語地名集(Gazetter)の構築

これまで英語、ゾンカ語の地名表記の標準化を進めるための試みがなかったわけではない。DDC は *Guide to Official Dzongkha Romanization* (1991 年)、*Samples for Geographical names of Bhutan in dzongkha and roman dzongkha with brief Guidelines* (1997 年) といったガイドラインを出版している。しかしこれらは少数の例、例えば県名・郡名を挙げているだけであり、その少数の例さえその後定着したとはいえ、本格的な地名集によって標準化を進めようという試みはいまだ実現していない。

こうした状況から、本研究ではゾンカ語地名表記について、情報通信省 (Ministry of Information and Communications: MoIC) が 2007 年に発行した地名コード表を基本資料としている。このリストは集落レベルまでの英語表記地名、ゾンカ語表記地名に自治体コードが割り当てられており、集落名の場合 4,294 件を収録している。1 件ごとに行政区画のレベルに対応するコードが振られているため、地名をデータベースとして扱いやすい。ただし英語、ゾンカ語の綴りには不備が多い。南部の現地地域言語地名とゾンカ語地名の両方がある場合、現地地域言語地名のゾンカ表記音写しか掲載されていないという、本研究の作業上致命的な問題もある。また、これ自体は地名とコードを対照したコード表に過ぎないので、地名と地図上の位置を対応できない。

地名の所在地の比定については NLC が 1960 年代にインド測量局 (Survey of India) の協力で発行した地形図やその改訂版 (1:50,000/2000 年、[図 6](#) 参照) をはじめ、2005 年の国勢調査 (Population and Housing Census of Bhutan 2005) の成果として国立統計局 (National Statistics Bureau: NSB) が公開している県ごとの主題図 (教育機関、医療機関の施設名、位置の確認に使用)、2008 年の総選挙の際に選挙区を示すために選挙管理委員会 (Election Commission of Bhutan: ECB、以下 ECB) が配付した郡別選挙区地図 (2011 年、[図 7](#) 参照) などを参照している。

これまで述べてきたように、これらの資料のそれぞれで同じ地名の表記が異なり、別名、別言語名が採用されていることがしばしばある。このうち ECB の選挙区地図は地図上の地名の一部にゾンカ語表記が添えられた唯一の例になる。また、NLC の地形図以外は単純に言えば政府関係施設の所在地が記載されているだけなので、一般的な汎用地図とは意味合いが異なる。唯一ゾンカ語表記のある ECB の選挙区地図は投票の便宜が目的なので、投票所のある集落の名は地図に必ず記載されるが、そうでない場所はリストで集落名が挙げられるのみで位置は確認できない。また、ECB の英語地名表記基準は、語源であるゾンカ語の表記に比較的忠実であるという利点がある一方で、政府やマスメディアが一般に用いている表記と大きく異なるローマ字化方式を採用している。

その後、2013 年に ECB が発行した選挙区関連地名表記の改訂リスト (英語表記版とゾンカ語表記版、さらにその対照表) や DDC によるゾンカ語地名表記正誤集 *Mingney Selwi Yoezer* (2019 年) を入手した。異同はあるものの、件数としては国立統計局のコードリストと同程度の地名を収録している。しかし、これらの資料が標準化の決定版であるかどうかは不明で、現時点でもメディアや政府機関によってはこの新基準を無視している。DDC の正誤表の場合、コードはおろか英語表記が対照されておらず、しかも収録順は英語表記地名のアルファベット順である。このため、検索しにくいというのに、綴りが変更された地名が明示されていないため、いちいち確認が必要という実用性を無視した構成になっている。

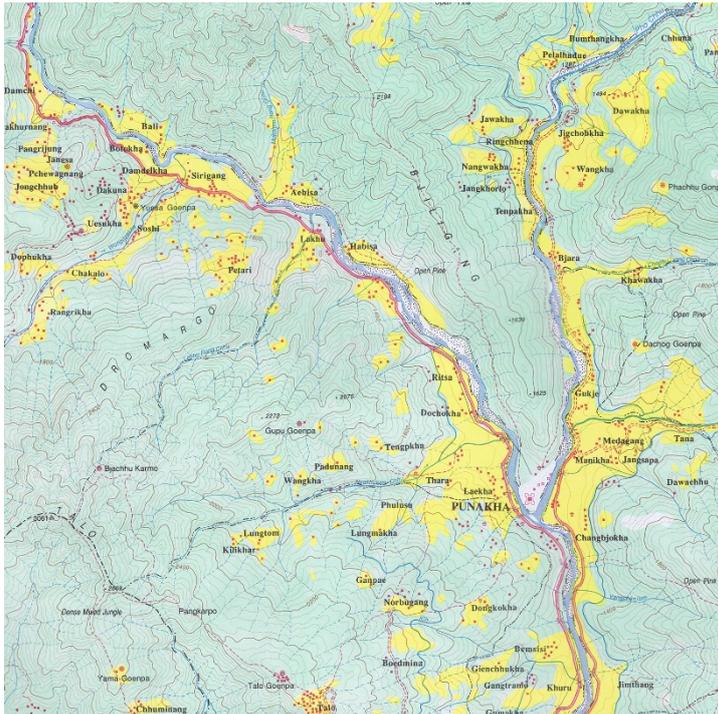


図 6 先行・比較事例

1:50,000 地形図 78E14 Punakha (インド・ブータン共同事業改訂版、2000 年)

ブータンの地形図は 1960 年代にインド測量局が作成し、2000 年前後にその 1/4 程度をインド・ブータン共同事業として改訂している。多くの区画が絶版になっており、国境付近は非公開なので全土をカバーしているとはいえない。また最新の改訂からすでに 20 年が経っており、自然地形以外の部分では実用に耐えない。

地名は集落レベルまでカバーしており、主要な寺院や歴史的建築物なども掲載されている。一方で調査が古いこともあり、学校や病院など近代的な施設はほとんど掲載されていない。

1960 年代の地形図の地名は、綴りだけでなく地名そのものが現在一般的なものと異なっている。また地名の発音の転写方法が恣意的であったり、英語的もしくはインド的であったりするため、語義やゾンカ語表記の字形を再現するのが難しい。

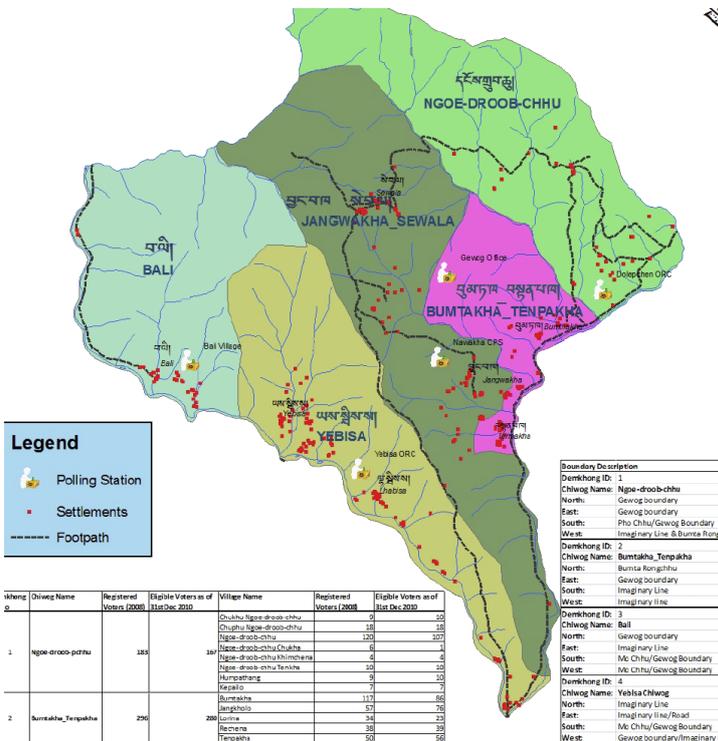


図 7 先行・比較事例

地方統一選挙郡別選挙区地図 (ECB、2011 年)

選挙区公示のために作成された地図で、地方議員の選出区である郡より小さな単位の行政区である庄 (Chiwog) で塗り分けられている。この庄名までゾンカ語表記が並記されていること、それ以下の村落にも一部ゾンカ語表記がそえられているという点で、一般的なものとしては唯一の「ゾンカ語地図」である。

ただし集落名のリストは英語表記のみであり、すべての集落が地図上にマッピングされているわけではない。投票所として学校が選ばれるケースが多いが、その場合も英語で学校名が記載される。

1 郡 1 枚で A4 に印刷されるため、郡の面積によって縮尺は変わる。集落名・行政区画名以外の地図要素としては川と道路が示されている。



(3) ゾンカ語地名の入力とマッピング

作業に用いた地名集はいずれもソースデータではなく PDF 形式の表として入手した。英語表記地名であれば PDF からテキストデータとして抽出するのは比較的容易だが、ゾンカ語地名は先述の文字コード体系のバリエーションや複雑な組版などの問題があるため、ゾンカ語ロケールを導入した環境でもデータ交換に文字化けが発生し、コピー&ペーストやテキスト抽出ができないものが多い。そのため、原則としてゾンカ語キーボードを使って再入力することになった。

また、学校名などの施設名は別資料をあたる必要がある。公刊されている学校名リストはすべて英語表記なので、地名がそのまま校名となっているケース以外ではゾンカ語表記は公式サイトや校章、表札などの画像データ、さらに教育省 (Ministry of Education: MoE) の教員配属辞令などの文書から採取した。また、正式な寺院名はゾンカ語表記ではなく古典チベット語表記が伝統的であり、プナカ県内のゾンカ語リストのみ内務文化省 (Ministry of Home and Cultural Affairs: MoHCA) の協力で入手することができた。

(4) 地図出力

等高線、河川、道路、地図記号といった地図要素はすべて Shape 形式で管理されており、文字装飾や配置は QGIS の「スタイル」として指定されている。最終的な印刷地図の整形、PDF 出力は QGIS の「コンポーザ」機能を利用し、レイアウトや全 76 枚のシートへの切り分け、全体地図、連続地図など、地図本体以外の要素も含めてすべて自動生成され、出力した PDF ファイルをそのまま地図として利用できる。このような地図印刷システム自体は“Gyamtscho Atlas Bhutan” 英語版のために開発したものを流用している。

なお、地図区画や ID は NLC の 1:50,000 地形図を踏襲しており、縮尺を 1:100,000 とすることで A3 サイズに収めた。これは既存地形図との比較、併用などの便宜をねらったものである。システム上は判型を大きくしたり、逆に同サイズで縮尺を大きくして枚数を増やしたりすることも可能だが、一般のカラープリンターによる出力など、現地事情を考慮した配付・利用の必要性からこのような仕様とした。

表 1 地名・施設名の収録数及びゾンカ語化達成率

種別	総数	収録数	ゾンカ語化	達成率
県 (Dzongkhag)	20	20	20	100%
郡 (Gewog)	205	205	205	100%
庄 (Chiwog)*	1044	—	—	100%
集落	約 5,000	2,516	2,333	93%
宗教施設	約 2,000	1,651	96	6%
教育施設	約 700	685	599	87%
医療施設	約 530	522	185	35%
合計		5,599	3,438	

注) 庄名は、原則として領域内の代表的な 1~3 の集落名を列記したものが用いられる。

試作にあたって地図そのものは可能な限りゾンカ語化した（表 1 参照）、その他の地図要素、例えばインデックスマップや地図ナンバー、シリーズ名、縮尺などは英語表記を残した。これはそこまでゾンカ語化してしまうと、今後こちらの保守作業が難しくなってしまうという理由による。また凡例などのゾンカ語化は現時点で終わっていない。

6. 研究成果

(1) 多言語地名テキストデータベースの構築

本システムではある地点とある地名が 1 対 1 で対応しているのではなく、ある地点に対して、複数の地名その他の情報が結び付けられている。地名の場合、情報通信省の英語表記、ゾンカ語表記、ECB の英語表記、ゾンカ語表記といった複数の表記、言語を収録した上で、最終的に地図印刷用に採用した表記を別に収録している。これは、標準化された正書法による綴りが確定できず、複数のソースを比較しながら作業を進める必要があったためである。しかし、このテキストデータベースは印刷地図作成の手段というだけでなく、さまざまな応用利用が可能である。

具体的な応用例としては、19 世紀から 20 世紀前半の英領インド政府関連の資料に出現する地名や、中世チベット人僧侶の伝記に出現する地名を現在の地名と結び付ける歴史地理研究ツールとして利用できる。また、英語地名をゾンカ語地名に変換できる変換システム用辞書も試作している。

(2) 成果物としてのゾンカ語表記地図の試作

現在に至るまで、最も収録地名数の多いブータン地図は NLC の 1:50,000 地形図シリーズ（インド・ブータン共同事業改訂版）である。本来なら 76 枚で全土をカバーすることになるが、現実的に使用できるのはそのうち 20 枚程度に過ぎず、それも発行年が古く、特に地名の信頼性が低い。近年発刊された農林省（Ministry of Agriculture and Forests: MoAF）の地図は収録地名が極端に少ない。ECB の選挙区地図は収録地名が多く、一部がゾンカ語表記になっており、また英語、ゾンカ語地名の綴りの信頼性が高いという利点があるが、用途が特殊であるため実用上の制約が大きい。例えば 1 枚の地図の区画が選挙区のアウトラインで切り抜かれることになるため、行政境界線上に道路や川があると橋や道路の反対側の地名が分からない。新聞に出てくる地名を地図上で確認するといった実用的な利用にある程度耐えられるのは、現時点において NLC の *Atlas of Land Cover and Institutional Facilities* 程度であろう。

本研究の成果物として試作したゾンカ語地図（図 8 参照）は、収録地名数、地理情報量の点でこれらの英語地図と同等以上の水準にある。

(3) 地図製作技術のローカライズに関する諸問題の検討

地図のローカライズは単純に表記言語の置き換えだけではなく、その地域の文化的背景が反映される。このため、地図記号においてもその国（文化）が色濃く反映されることになる。ブータンの場合一部の県のゾンは、行政的には県庁であると同時に宗教的には僧院であり、歴史的には国宝級の文化遺産である。一方で単純に近代的な県庁としての機能し

かもたないゾン、現在は遺跡であるゾンもある。これを地図記号でどう表現するかはなかなか悩ましい問題である。

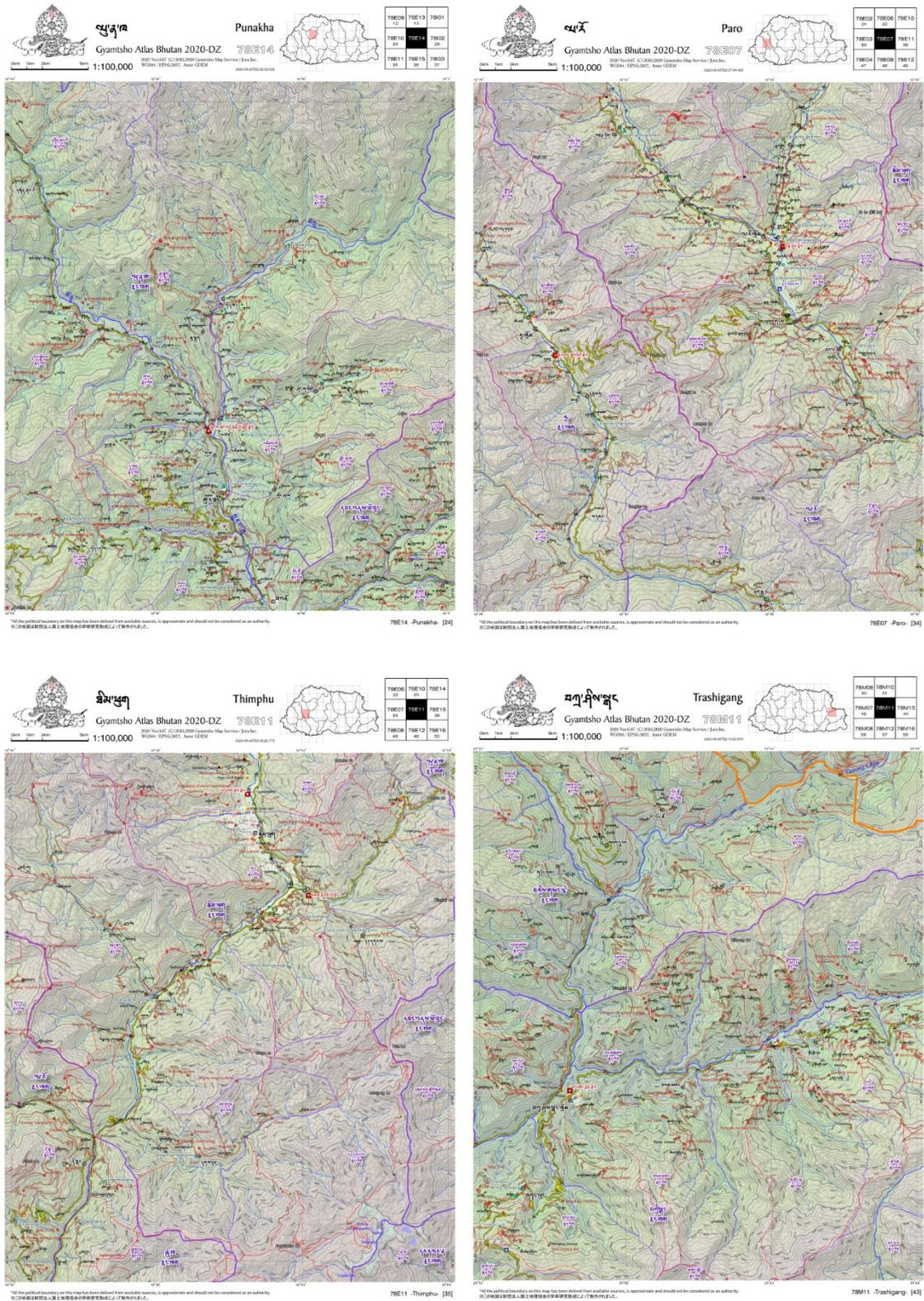


図 8 本研究で試作したゾンカ語表記地図帳(全 76 枚)の一部

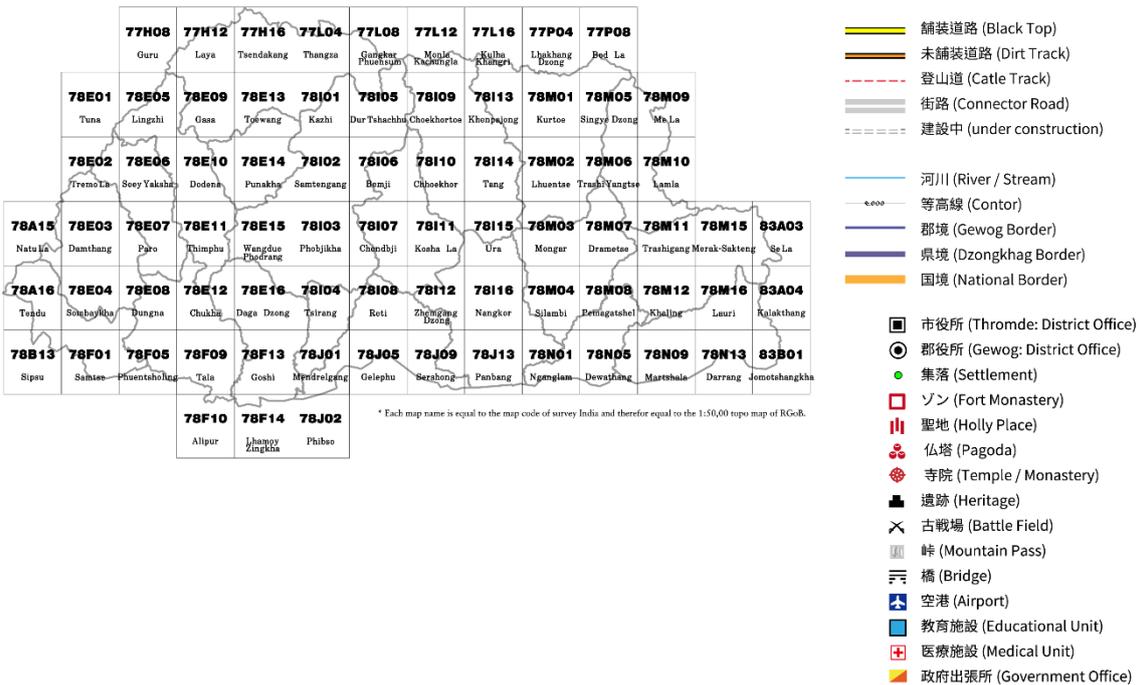


図 9 作成したインデックス及び「ブータン風地図記号」凡例

ブータンには近代化以前の地図文化がまったく存在せず、また、近代化以降の大規模な地図製作プロジェクトは例外なく道路建設、土地開発、技術移転といった海外諸国・機関による経済的・技術的支援と連動したものであった。このため既存の地図は単に表記言語として英語を採用しているというだけでなく、あらゆる面で欧米の地図文化をそのまま受け入れたものとなっている。したがって、例えば郵便局や銀行のような地図記号もローカライズされていない。NLC の地図が採用している地図記号はほとんどが幾何学図形の色分けであり、シンボル化された記号は飛行場（飛行機）、発電所（稲妻）、病院（赤十字）の 3 つにすぎない。本研究では仏塔、峠、橋など伝統文化関連の「ブータン風地図記号」の試作も行った（図 9 参照）。

(4) 社会言語学

ゾンカ語地図の製作上の選択肢として、例えば標高や経度緯度の数値をアラビア数字とするかチベット数字とするかといった問題がある。本研究にはゾンカ語化のトライアルとしての意味もあるため、これらにもゾンカ文字を採用した。その際、数表のようなものでもゾンカ語文書ではチベット数字を使っていることを確認している。

ゾンカ語化を進める中で、英語表記では「村落名+A」「村落名+B」、ゾンカ語表記では「村落名+ᄀ」「村落名+ᄁ」となっている 2 つの村落があることが分かった。ᄀ、ᄁ は英語の ABC、日本語のイロハに相当するゾンカ文字なので意味としてはおかしくないが、本来のゾンカ語の命名法ならこの 2 つの村落は川の上流下流、あるいは標高の高低によって「村落名+上」「村落名+下」と呼び分けるのが通例である。ゾンカ語以外の言語の同名（あるいはゾンカ語では同音）の 2 つの地名を区別して表記するための苦肉の策で、2 つの村落

が実際にこの通り呼び分けられているわけではないだろう。

ゾンカ語やその他の地域言語は言語学分野での先行研究があるものの、多言語社会の実状や近代化の影響といった社会言語学的な研究は少ない。ゾンカ語地名の収集やゾンカ語地図の作成はその新しい切り口となり得る。特に外国人研究者が足を踏み入れることがほとんどなかった南部地域における言語状況の解明の手がかりになることが示唆された。

(5) 言語政策・文化政策の基礎資料の収集

本研究の作業を進めるにあたって、先行事例（地図）の収集及び検討を網羅的に行ったが、同時に地名収集のために各種のリストを探索することとなった。既に述べたように大規模なゾンカ語地名集として情報通信省、DDC、NLC のリストを参照しており、学校名では教員配属辞令、寺院名では内務文化省の文化財リストを利用している。ブータンには地名辞典が存在せず、また、政府内公文書や国際プロジェクトの報告書などの資料も基本的に英語で記述される。英語－ゾンカ語辞書は DDC のものをはじめいくつかあるが、これらは基本的に対訳辞書であり百科事典的なものはないため、まとまったゾンカ語の固有名詞リストの入手が困難である。本研究の成果として、これらのゾンカ語地名、施設名、機関名のリストの所在やその有無がある程度確認された。

7. 考察

研究に着手した時点では、従来の英語地図の地名表記に実用上問題になるほどの表記の揺れがある原因はゾンカ語のローマ字化正書法が浸透していないためであり、ゾンカ語地名の揺れはそれに比べて問題にならないほど小さいと考えていた。そのため、入力や組版といった技術的な問題を解決できれば、綴りや地名の揺れの問題はむしろ解決されると想定していた。具体的にはゾンカ語表記については DDC の協力を仰ぎ、そのガイドラインに沿うことによって標準化・統一化できると考え、現地を訪問して NLC の担当部門に地図製作面での協力を、DDC の担当部門にゾンカ語表記についての協力を打診し、基本的な了解を得た。

しかしながら、現実にはゾンカ語の地名表記についてもいくつかの未解決の問題、あるいは政府内での意見の対立があり、オフィシャルなゾンカ語地名表記の確定という最も基本的な作業でつまづくこととなった。言語政策は政治的な意味も大きいため、政権の交代によって方向が左右されるといった事情もあった。ゾンカ語はブータンにとって象徴的に重要な文化要素ではあるが、実用的には、少なくとも書き言葉としては英語に比べて重要視されない。一方、地図は本来なら地域の言語や文化と密接に結び付いた分野のはずだが、開発や国際協力の分野でニーズがあるため、その観点での利点、つまり英語表記の必要性が高い。ゾンカ語の専門機関である DDC や文部行政を司り学校教育でゾンカ語教育を行っている教育省、土地管理を行っている部門や地理情報や地図を管理している部門でもなく、選挙管理委員会が地名表記の標準化のイニシアチブを握るといった状況は、ブータンの社会的背景を反映したものと思われる。

一般的な基準でいえば多言語・多民族国家であるブータンにおいて、ゾンカ語はナショナル・アイデンティティの中核となる要素のひとつとして位置づけられている。それにも

かわらず、その国語で表記された一定の水準の国土地図が存在しないというのはいかかなものかというのが、本研究の出発点であった。このことは、社会的なニーズという面からいえば英語表記地図の必要性、便宜性が高いという現状を了解したうえで、別の選択肢の可能性を模索するものであったともいえる。実製作を行った結果、ゾンカ語表記地図にもそれなりの利便性があり、一定のニーズがあるのではないかと考えるようになった。例えば母語がゾンカ語である地域の近代学校教育を受ける機会のなかった人たちや僧侶である。教授言語は英語であり、ゾンカ語には発達した敬語表現や慣用があるため、学校教育を受けた人たちにとっては英語の読み書きのほうがゾンカ語よりも得意であることが一般的である。しかし、ゾンカ語の記法自体は基本的にはチベット文字を利用した表音なので、例えば地名のような固有名詞の発音を表記する方法が、アルファベット表記（ローマ字表記）より難解だとは一概にいえない。

一方でブータンのような多言語社会において、英語版、ゾンカ語版地図といった言語で分けた地図製作が必然的なものであるかについて改めて考えるようになった。学校教育を受けていない人も利用者として想定していると思われる ECB の選挙区地図は英語、ゾンカ語の地名 2 言語表記が一部取り入れられている。このような 2 言語表記や、あるいは山名や川名、集落名はゾンカ語、学校や病院のような近代的施設は英語、寺院名は古典チベット語で表記するといったハイブリッド地図も、現代ブータンの重層的な言語事情からいえば意外に利便性が高いかもしれない。

このような二重性は、必ずしも表記言語や文字の違いだけに由来するとは限らない。例えば、現代的な地理認識の中で重要なランドマークとなるのは自動車道路の分岐点といった交通の要衝、その地域の役場や学校、病院がある場所になりがちである。しかし、自動車道路の分岐点は実際には所在する場所、あるいは所在地の集落名ではなく、近隣の大きな村の名前や広域の地域名によって呼ばれることが多い。結果的にその地名の由来である集落は我々が土地名として認識している場所とかなり離れているケースがしばしばある。現在「××峠」といえばたいいてい人は自動車道路の最高地点だと認識するが、歴史的な峠や、そこから名づけられた「××峠村」は別の場所にある。これは地理認識の時間的変化だといえるが、問題はブータンの場合このような変化が非常に短期間で起きていることにある。例えば道路網が毎年倍増する勢いであることを考えれば、10 年後、20 年後には更に激しい変化が予想される。今回のゾンカ語化作業でゾンカ語表記が確認できない未解決地名にはそうした「新地名」がかなり含まれている。地名データベースの整備やゾンカ語地図の製作は、このような社会変化によって失われていく可能性の高い、国土の歴史・地域文化の伝統の記録・保存のうえでも重要であろう。

8. まとめと課題

地図製作上圧倒的に大きな課題となったのは、繰り返し述べてきたように、標準的なゾンカ語表記が確立されておらず、地名集やそれに代替することができるリストが整備されていないという点である。こうした文化的・社会的条件を除外すれば、Unicode によるテキスト処理の環境が整備されてきたことで、ゾンカ語処理の課題の多くは解決済みだといえる。

地名に関しては90%以上をゾンカ語表記化し、現在一般に入手可能な最も詳しい英語版全国地図とほぼ同等の収録数になっている。これら近年発行された詳細地図はいずれもブータン政府の地理情報システム（GIS）のデータをもとにしている。そのため、道路、学校、病院、役所といった公共施設、また森林、耕地、河川といった資源・生産関係についての情報が網羅されているが、逆にいえば寺院、私有の産業施設、観光地、放牧地、温泉、ホテル、ショッピングセンターといった政府が直接関与しない、いわば民間セクターの情報は基本的に含まれていない。日常的な地理感覚ではこちらがランドマークとして重要なケースも多く、またNLCの地図帳の場合、地図記号が記載されるだけで施設名は掲載されていない。本研究ではブータン人の日常生活に密着し、伝統文化を意識した地図を目指すという意図から、それらを可能な限り収録しようと試みている。しかしながらプナカ県以外では利用可能な寺院名のゾンカ語表記のまとまったリストが入手できなかったため、寺院名を英語表記のままとせざるをえなかった。

また、ゾンカ語化済みの地名についても、作業の前後関係や収録元の関係から、ECBやDDCの発表している新表記基準に合致しているかの検証や、それらの新表記基準に含まれない地名・施設名の表記との整合性のチェックなどが今後の課題として残されている。山名、河川名、峠名など、おそらくまとまったゾンカ語表記によるリストは存在しないが、ランドマークとしては重要な地名をどのように処理するかについても、決定的な手法を確立できなかった。

それ以外の今後の課題として大きなものは、フォントのデザインとバリエーションの問題である。日本語でも英語でも通常、出版物の文字サイズは手書き文字よりもかなり小さく、またフォントデザイン自体が、そのような小さな文字でも可読性が確保されるように工夫されている。しかしチベット文字は手書き文字をそのまま写し、手書きの場合と同じ文字サイズで印刷することが基本になっており、その意味で印刷用書体というものは整備されていない。地図では文字配置上の制約が多いため、一般的な印刷物に比べても小さな文字サイズを使用したいケースが多いが、ウチェン体はそういった用途に向いておらず、ロ・イク体に関してはそれよりもさらに向いていないといえる。

また、母音記号や結合文字の関係から、文字高にかなりバリエーションがあるため、コンパクトに収めにくいだけでなく、下線や文字囲などの文字装飾が使えない。同じ書体の太さによるバリエーション（ウェイト）という概念もない。スペースとの関係でいえば、アルファベットであればコンデンス体のように文字幅を抑えた書体、イタリック体のような強調書体を使い分けることができるが、それもできない。文字色と文字サイズ以外で文字の表現にバリエーションをもたせることができないことはデザインだけではなく、可読性の面からも大きな制約となっている。また、ブラーフミー系文字に共通する性格だと思われるが、文字列の基準となる仮想的な線がデザイン的にはっきりしている（特にウチェン体）ため、直線状に配置した場合にはまとまりがよく、地名をまとめたデザイン要素として表現するには向いているが、例えば川に沿って曲線状に配置するといった手法に向かない。

テキスト処理の面では、地名50音ソートの手法と基準を詳細に検討することができなかった。今回の成果物は出力地図だけとし、地名索引は用意しなかった。これはゾンカ語表記が確定できない地名が歯抜け状に残っているという状況や、表記基準の標準化そのも

のを完全に進めることができないという事情によるものだが、仮にそれができたとしてもゾンカ語地名のソートにはノウハウが必要であると思われる。

これらの課題を解決するためには、ゾンカ語のオーソリティーを含めた専門家、各分野の技術者の協力や助言が必要であるとともに、成果物であるゾンカ語地図に対するブータン人の評価を広く確認する必要がある。今回試作したゾンカ語地図をサンプルとして、**DDC** などゾンカ語関係者、**NLC** など地図製作関係者の評価を仰ぎ、また **Web** 地図として一般公開することでブータン人一般の意見を集めることを次の目標としている。また、地図製作の過程で得たゾンカ語の言語状況に関する知見などは、研究代表者・共同研究者が所属する各学会において発表、論文化する予定である。

ゾンカ語綴りの確認のために参照した主な資料

発行元	発行年	資料名	収録数	英語	ゾンカ語	表記	配列方法	備考
DDC	1991	Guide to Official Dzongkha Romanization	約 100	○	○	旧	全国／首都周辺	サンプルとしてゾンカ語表記とローマ字化の例が収録されている。現在使われていないアクセント記号を併用したローマ字化方式を採用している。
DDC	1997	Samples for Geographical names of Bhutan in dzongkha and roman dzongkha with brief Guidelines	約 200	○	○	旧	県／郡	県名・郡名。
情報通信省	2007	Data Standards (仮称)	約 4,000	○	○	旧	県／郡／集落	構成は行政区画レベルごとの別表だが、コードによって階層化されている。
ECB	2013	English Name Change Order	約 5,000	○		新旧	県一郡一庄	それまで利用していた選挙区名の綴りを改訂した正誤表。
ECB	2013	Dzongkha Name Change Order	約 5,000		○	新旧	県一郡一庄	
ECB	2016 (?)	Names of Villages Under 1044 Chiwogs	約 5,000	○	○	新	県一郡一庄一集落	英語・ゾンカ語版の正誤表を統合し、同じ表記ルールを村名にも当てはめたものと思われる。
ECB	2016	Atlas of the Local Government Demkhongs 2016 (英語版)	約 5,000	○		新	県一郡一庄一集落	配列はどちらもアルファベット順。選挙区地図中の付表だけをまとめたものと思われる。
ECB	2016	Atlas of the Local Government Demkhongs 2016 (ゾンカ語版)	約 5,000		○	新	県一郡一庄一集落	
DDC	2019	Mingney Selwi Yoezer	約 5,000		○	新旧	県一郡一集落	配列は県が西から東へ、郡・集落がアルファベット順になっている。ECB の改訂綴りと異同があるかどうかは未検証。

注) ブータンの行政区画は、上のレベルから Dzongkhag (県) — Gewog (郡) — Chiwog (庄: 地方選挙区) — 集落 (村) という 4 段階になっている。