様式 C-19

# 科学研究費補助金研究成果報告書

平成２３年３月３０日現在

機関番号：１７１０２
研究種目：基盤研究（C）
研究期間：2007～2010
課題番号：１９５２０３５１
研究課題名（和文）　中国西南部の民族文字の字形集合に関する情報理論的研究

研究課題名（英文）　A Study on the Script Sets of Minority Peoples in Southwestern China

研究代表者
鹿島　英一（KASHIMA EIICHI）
九州大学・留学生センター・教授
研究者番号：２０２５３７００

研究成果の概要（和文）：
中国西南部の(雲南省と貴州省に跨る)雲貴高原やその周辺地域に分布する非漢字系の象形文字(絵的文字)や(チベット文字やラテン文字とは関係の薄い)固有の音節文字を対象にして、情報理論的解析を試みて、定量的な面からその特徴を把握することに努めた。具体的には、納西(ナシ)象形文字(東巴文字)、納西標音文字(哥巴文字)、規範彝文、傈僳音節文字などの外に、水文字も対象とした結果、以下のことが判明した。即ち、１．規範彝文、東巴文字(及び哥巴文字)と傈僳音節文字は明らかに漢字とは別系統に属し、これに水書の文字を加えて、一つの文字圏と見なすことに特に問題が無い。２．音標文字や漢字といった直線状字体の研究で得た文字集合の特徴を定量的に把握するための諸指標(各文字の文字単位数の分布状況、各文字と抽出した文字単位集合の対称型の分布状況など)が、象形文字や固有の音節文字といった曲線状字体に適用した場合でも一定の結果が得られて有効であること。また、論展開の必要性に基づいて女真文字(北方系擬似漢字)の字形集合について調査した。

研究成果の概要（英文）：
The proper hieroglyphic syllabaries, popularly unknown, in shape and origin are scattered over the south-western highlands of China. In this Research, I have investigated and analyzed, concretely, characteristics of these script-sets. Consequently, followings below are clear. ①A script group of modern Yi, Nasi Donba and Geba, secret Shui, Lisu syllabary and etc are unique. ②The way of analyzing I have created ago are effective as well in picturesque hieroglyphicals as in straight-line scripts (as alpabeticals and modern (Japanese,Chinese and Korean) Kanji-s.
In Supplementary, I also investigated the Jurchen script-set as northern type Square-script.

交付決定額

（金額単位：円）

| | 直接経費 | 間接経費 | 合　計 |
|---|---|---|---|
| ２００７年度 | １，１００，０００ | ３３０，０００ | １，４３０，０００ |
| ２００８年度 | ８００，０００ | ２４０，０００ | １，０４０，０００ |
| ２００９年度 | ８００，０００ | ２４０，０００ | １，０４０，０００ |
| ２０１０年度 | ８００，０００ | ２４０，０００ | １，０４０，０００ |
| 年度 | | | |
| 総　計 | ３，５００，０００ | １，０５０，０００ | ４，５５０，０００ |

研究分野：文字学
科研費の分科・細目：言語学
キーワード：中国語学、中国の民族文字論

## １．研究開始当初の背景

研究代表者は、情報処理革命の進む近未来の世界で必須とされるにも拘わらず、国内外に殆ど先行研究を見ない分野に於いて、二十年程にわたり全くの第一歩から考察を積み重ねてきたが、それは主に元来が通信(情報)工学の出身であったという個人的な経歴に依っていた。

始まりは「文字の統計学的処理」(博士論文、平成５年)で、日頃、文化・文明論的な議論の対象になりがちな文字を、対象をラテン系文字を始めとする音標文字に絞って、情報通信符号の側面から考察したこと(音標文字の情報理論)である。代数幾何学、集合論、推測統計学などの数学的手法を駆使したその成果を踏まえ、次に研究課題を漢字(漢字の集合論)に移した。現在と近未来の世界での普遍的な文字技術の確立に、表意・表語文字からの貢献が欠かせないからである。

その結果、例えば日本(常用漢字)やシンガポール周辺地区・中国(簡体字)や台湾・香港澳門・韓国(繁体字)から東南アジア・北米・オセアニアの新僑民地区へと繋がる新「漢字圏」の漢字の諸集合も、各字形を構成する「文字単位」(字形の基本単位：検字用の部首など)のレベルでは音標文字と同様の手法が適用でき、かつその集合の具体的な実相が統計的・定量的な面から確実に把握できるようになった。つまり、日本(新字体、即ち常用漢字:1945 字)、中国(簡体字:「常用」2418 字)、韓国(旧字体:常用字1800 字)の三種類の漢字(字形集合)を扱った結果、(相互間の差が最大618 字あるにも拘わらず)どの文字単位集合も総数がほぼ同じ(六百余)な上、その(幾何学的)「対称型」分布もほぼ同じことが判明している。

また、日韓中の漢字自体についても相互間の共通字形率、「字形の(空間)配置」型の分布、各文字をなす文字単位数(／文字)の分布の調査の結果として、共通字形率(日韓で 3 割強と1/3 弱、日中で 5 割弱と 4 割弱)は三集合の差異が例えば(ギリシャ系音標文字の)ラテン大文字とギリシャ大文字とキリール大文字の関係(かそれ以上)に相当することが判明し、優に一千年を越えて存続した(ベトナムや海外華僑域を含む)旧漢字圏に於ける二十世紀の文字改革が担った別の面が明らかとなったし、(基本的には類似する)後の二者からは相互間の相違点も明らかと成った。(対称型と空間配置(型)は共に曾て研究代表者が案出し、有効性を確認した字形の特徴を示す指標である。)尚、以上は「音標文字の情報理論」と「漢字の集合論」から成る『文字の情報理論序説』(拙著:1997 年:風間書房)、日中韓の漢字(字形集合)を扱った『漢字の情報理論』(拙著:2006 年:白沙ヶ濱)として、既に刊行している。

その後、引き続いて擬似漢字を主対象とする研究に着手した。(漢字と擬似漢字に共通する)漢字系文字の特徴を、定量的な面から把握するためである。方法論は基本的に漢字に準じたものであり、「擬似漢字」とは具体的には契丹文字、女真文字、西夏文字を指している。(程度は異なるが)何れも解読課程の途上にあると言える。同類で、(日本の国字(「粂」「峠」等)の類が豊富な)方塊壮字、字喃、方塊白字および同系の「派生漢字」(侗文字、布衣字)は文字資料の入手難など技術上の問題はあるが、集合規模が大きい方塊壮字や字喃も(字数が約五千字の台湾「常用国字」で)漢字で確立した方法が有効なようで、操作自体は煩雑だが、本質的な問題は多くない。以上が本研究の学術的背景である。

## ２．研究の目的

本研究の目的は、中国西南部の(雲南省と貴州省に跨る)雲貴高原やその周辺地域に分布する非漢字系の象形文字(絵的文字)や(チベット文字やラテン文字とは関係の薄い)固有の音節文字を対象にして、情報理論的解析を試みて、定量的な面からその特徴を把握することである。具体的には、納西(ﾅｼ)象形文字(東巴文字)、納西標音文字(哥巴文字)、規範彝文、傈僳音節文字などの外に、水文字も対象としたい。これらと漢字圏の間に曾て交流があったことは確かなようだが、大抵は明らかに異なる文字圏を構成しているからである。現時点で、現存する唯一の象形文字として珍奇な目で見られがちの東巴文字、その先祖が源甲骨文字であるとの説が取り沙汰され始めた彝文字の研究は、近年になって地中から姿を現し始めた(エーゲ文明、マヤ文明、アンデス文明などに匹敵するかもしれない)長江文明と巴蜀(共に四川省低地)の文字の研究に新しいアプローチを開く可能性を秘めており、同時に本研究の特徴でもある。

研究の分析視点とそれに対応した研究課題は具体的には以下のように纏められる。

第一は、従前の研究で得た文字集合の特徴を定量的に把握するための諸指標が、象形文字や固有の音節文字にも適用できるかを判断することである。というのは、既述の事柄の研究対象は直線状字体であったからである。特に、字形の対称型の有効性はそれを前提としていた。少なくとも東巴文字は絵的(曲線的)であり、哥巴文字も古い彝文字もある程度(古代エジプトの神官文字ほど)は曲線的である。他では、字形の配置(型)などは工夫がいるかもしれないが、文字単位数(／文字)は余り問題はないだろう。何れにせよ、全くの直

線(と円)状である規範彝文での検証も含め、(必要で可能なら新指標を導入して)有効な諸指標を確定するとこが必要である。

第二は、東巴文字、哥巴文字、規範彝文、字、傈僳音節文字文字、水文字などの文字字形集合を対象に、第一の課題で得られる諸指標を使って、定量的に調査することと、その結果を従前の研究で得た成果と比較分析して、(可能な範囲で)これらの非漢字系文字の特徴を定量的に解析把握することである。これらは文字数が多い東巴文字でも一千数百字であり、字形集合の規模の上でも可能であろう。そして、最後に子の時点までで得られている統計データを全て用い、東アジア地域の固有文字群の実相を明らかにしたい。

３．研究の方法

初年度は基本的に規範彝文を対象にして、字形集合の特徴を定量的に把握するように努めた。無論、雲南省との境界地域にある四川省の涼山地方(方言分化の大きい彝語の標準音地域)の西昌市に直接出掛けて、象形文字に由来するその古文字や規範彝文の資料を調査し、多様な入手を試みて資料を充実させた。その結果、『彝語大詞典』等を基に、対象となる文字集合(1164 字)を確定し、それに対し情報理論的解析を試みた。

文字集合には 26 種の部首(『彝文検字本』)があり、(拼音方案のローマ字表記を分類の補助に使いつつ)基本的には曽て研究代表者が始めた漢字での方法論に倣って、これを独立の文字単位と(そうでない)小図形に仕分けすることから始め、各文字をなす文字単数数の分布、各文字単位字形の(幾何学的)対称類型分布、各文字の配置類型分布の一覧表を得た。また、同時に各文字の対称類型分布の一覧表も得た。これは、字数が(約２倍と)近い漢字(常用字)と大きく異なり、字数が(1/15～1/30 と)小規模の音標文字の文字集合に似たデータであったが、説明がある程度つくものである。

この他、本年度は水文字傈僳文字の資料収集に雲南省や貴州省方面にも出掛け、一定の成果を得た。

二年目は納西象形(東巴)文を主対象にして、資料調査と解析を実施した。当然だが、雲南省の麗江地区(納西語の標準音地域)に行って資料収集をすると同時に、その周辺に分布する(東巴文字に)由来する納西標音(哥巴)文字の資料も充実させた。

また、これらと『納西象形文字譜』と基に、対象となる文字集合(1340 字)を確定し、古漢字(甲骨文)にも似たこの文字に対して、情報理論的解析を試みた。具体的には、古漢字の『説文解字』や規範彝文の『彝文検字本』に代替できると見た諸資料を用いて、基本的には曾て研究代表者が始めた(現代)漢字で執った方法論に倣って、これを独立の文字単位と(そうでない)小図形に仕分けすることから始め、文字単位集合を抽出した。また、各文字をなす文字単位数の分布、各文字単位図形の(幾何学的)対称類型分布の一覧表を得た。

結局、東巴文字の文字単位集合は構成素数や対称類型分布で(字形構成などで似た現代漢字よりは)規範彝文に近いとデータを示すことが判明した。

また、本年度は更なる対比データを得るために、西南部と同時期に中華文明の北方方面に曾て分布した女真文字に対しても同様の集合論を試みて、文字集合の統計データを得た。この試みもこれまで実施されたことがないものである。

この他、本年度は水文字や類似の民族文字の資料収集に中国の当該地域方面に出掛け、一定の成果を得た。

三年目の平成 21 年度は雲貴高原東南麓域(貴州省南部)の三都水族自治県を中心に分布する水(スイ)族の伝統文字、即ち水書(白書と秘伝の黒書)を主対象にして、資料収集と解析を実施した。

尚、水書は 1980 年代(新中国)になって研究の途についたばかりで、経典資料は近年やっと刊行されるようになった。(字形の他に)読音も方言に依って違う状況だが、今回は『水書常用字典』の常用字(471 字)を対象とした。具体的には、古漢字風と東巴字風の両種が混在するこの字形集合に対して、基本的には曾て研究代表者が始めた(現代)漢字での方法論に倣って、これを独立の文字単位と(そうでない)小図形に仕分けすることから始め、文字単位を抽出した。また、各文字をなす文字単位数の分布、各文字単位図形の(幾何学的)対称類型分布の一覧表を得た。また、同時に各文字の対称類型分布の一覧表も得た。

結局、全てが単体字のこの文字集合のデータ(定量的特徴)は漢字よりは東巴字(特に、890 字の単体字集合)と類似点が多い(規範彝文とは概して似ていない)ことが判明した。

また、この他に本年度は傈僳(リス)音節文字の資料収集に加えて、代表的な擬似漢字である越南の字喃(チューノム)の予行的解析も実施した。

最終年度は高原北部の(雲南省)維西傈僳族自治県を中心に分布する人口が60万人程の傈僳(ﾘｽ)族の間に、二十世紀初頭に突如として出現した固有の音節文字を主対象にして、資料収集と解析を実施した。尚、この文字は同時

期にキリスト教布教用に現れたﾌﾚｰｻﾞｰ系の老倮傈文や滇東北部の苗系の禄勤倮傈文とは異なる。

具体的には、創字者の凹士波(Ngual sseixbbo,漢人式名:汪忍波)の残した「識字課本」にある1286字を対象(の字形集合)とした。集合は漢字風や納西族の哥巴(ｹﾞﾊﾟ)文字風(漢字や哥巴文字や東巴文字を借源字にして変形した字形)と独自形が混在するもので、基本的には曾て研究代表者が(現代)漢字で開始し、規範彝文、納西東巴(ﾄﾝﾊﾟ)文字、水書などで展開した方法論に倣って、集合論を展開した。

先ずは、手書きで難解な異体字を同定する過程で、1286字を936字に絞った。また、この集合に関し(幾何学的)対称類型分布の一覧表を得た。結果は規範彝文(1164字)や東巴文字(1339字)とほぼ同じ傾向を示すことが判明した。即ち、左右対称(Ⅱ型)と非対称(Ⅷ型)で合わせて約九割(余)、またその内のⅡ型が約三割前後、残りの対称型は合わせて約一割である。(但し、規範彝文は途中で字形が90°回転した事情が割り引いてある。)資料公開が近年で、研究書が殆ど無いため、936字の大半を占めると見える単体字から、複合字を厳密に分離することは今回は差し控えたが、前年報告の水書(467字中の464字が単体字)と同じ傾向の文字として理解できる。

４．研究成果

規範彝文は古漢字(甲骨文字)の源字の可能性もある各地の彝文字(とその90°回転形)を利用した文字だから当然だが、東巴文字(及び哥巴文字)と倮傈音節文字も明らかに漢字とは別系統に属すことが改めて確認できた。(無論、この三種の文字集合間には創字の時期が実質的に一千年近い開きがあるため、(漢字圏との接触の結果)一部に漢字からの派生字があるが、本質的な事柄ではない。)

(異体字を除いた)文字数はどれも一千字前後で、これに実体の明らかになり始めて日の浅い水書の文字を加えて、一つの文字圏と見なすことに特に問題が無いことが本研究の成果の第一である。尚、根拠は(個々の文字に当たれば自ずと知る)文字の構成法や類似感覚に加えて、研究目的で述べたことの二つ目、即ち各文字集合の統計データに現れた傾向の類似性である。

第二は、研究目的で述べたことの最初のもの、即ち音標文字や漢字といった直線状字体の研究で得た文字集合の特徴を定量的に把握するための諸指標が、象形文字や固有の音節文字といった曲線状字体に適用した場合でも一定の結果(３.＜研究の方法＞の末尾で記したことなど)が得られることが判明したことである。

５．主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計５件) 全て単著

①鹿島英一「規範彝文の集合論」『東北大学言語学論集』、査読有、第16号、pp.43-68.2007年

②鹿島英一「納西東巴文字の集合論」『東北大学言語学論集』、査読有、第17号、pp.39-68.2008年

③鹿島英一「女真文字の集合論」『地域文化研究』、査読有、第７号、pp.45-98.2009年

④鹿島英一「水書の集合論」『東北大学言語学論集』、査読有、第18号、pp.21-44.2009年

⑤鹿島英一「倮傈音節文字の集合論」『地域文化研究』、査読有、第９号、pp.67-90.20011年

〔学会発表〕(計１件)

① 鹿島英一「倮傈音節文字の集合論」地域文化研究学会、(2010.12.26)、於福岡市

６．研究組織(1名)

(1)研究代表者

鹿島　英一 (KASHIMA　EIICHI)

九州大学・留学生センター・教授

研究者番号：２０２５３７００

(2)研究分担者

なし　　(　　)

(3)連携研究者

なし　(　　)