

日本語の系統とその遺伝子的背景

松本克己

2012年11月15日 千葉大学文学部

1 はじめに

本日の講演の題目に掲げられた「日本語の系統」という問題は、今から百年以上も前から内外の大勢の学者が取り組んできて、未だに決着のつかない難問とされてきました。通常、言語間の同系関係というものは、人間の親族関係と同じように、同じ祖先を共有するというような意味で、その祖先に当たる言語を「祖語」、同じ祖語を共有する同系諸言語の全体を「語族」と呼びます。

現在、日本列島を含むこのユーラシア大陸には、数にして2千以上の言語が話されていますが、そのほとんどの言語は、このような同系関係によって10個余りの語族の中に纏められています。そして、これら語族内部の言語間の同系関係は、それぞれの言語の基礎的な語彙、例えば、1から10までの基本数詞、目・鼻・手・足のような身体名称、あるいは身近な親族名称などを比較すれば、簡単に確認できるような性格のものです。

しかし、日本語の場合、このような基礎語彙のレベルで、同系関係が確かめられるというような言語はこれまで全く見つかっていません。この意味で、従来の歴史・比較言語学の立場からは、日本語は外部に確実な同系言語を持たない、つまり系統的に孤立した言語として位置づけられてきました。現在、ユーラシアにはそのような系統的孤立言語が10個近く数えられるのですが、実は、その中の半数近くがこの日本列島とその周辺に集中しています。すなわち、日本語のほかに、列島北部のアイヌ語、その対岸のアムール下流域と樺太で話されているギリヤーク（別名ニヴフ）語、そして朝鮮半島の朝鮮語も、日本語と同じように系統不明の言語とされています。つまりこれらの言語は、見方を変えれば、それ自体が単独でひとつの語族を形成するとも言えるわけです。

従って、このような言語の系統関係を明らかにするためには、伝統的な歴史・比較言語学の手法とは違った何か別のアプローチを試みる必要があります。従来の歴史言語学で「比較方法」と呼ばれる手法は、主に形態素や語彙レベルの類似性に基づいて言語間の同系性を明らかにしようとするものですが、このような手法でたどれる言語史の年代幅は、大体、5～6千年あたりがその限度と見られています。日本語やアイヌ語の系統が不明だということは、結局のところ、これらの言語の系統関係が従来の比較言語学では手の届かない遠い過去にまで遡るという意味にほかなりません。

2 類型地理論から探る言語の遠い親族関係

それでは、このような語族という枠を超えた“言語の遠い親族関係”を探るためには、どのような方法をとったらよいのか。日本語の系統問題に取り組むに当たって、まず最初にぶつかるのは、このような方法論上の問題と言ってよいでしょう。そして私の場合、あれこれ試行錯誤の末にたどりついたのが「言語類型地理論」と呼ばれるような手法です。

これは簡単に言うと、それぞれの言語の最も基本的な骨格を形作ると見られるような言語の内奥に潜む特質、通常は「類型的特徴」と呼ばれるものですが、そのような言語特質を選び出し、それらの地理的な分布を通して、それも世界言語の全体を視野に入れた巨視的な立場から、それぞれの言語または言語群の位置づけを見極めようとするものです。このような目的のために、特に日本語に焦点を据えて、私が選び出した類型的特徴とはどのようなものか、またそこからユーラシアの諸言語の中で日本語はどのように位置づけられるのか、それを結論的に纏めてみると、配付資料の「表1」<類型地理論的考察から導かれたユーラシア諸言語の系統分類>というような形となります。

そして、この表の下段に挙げた8つの項目が最終的に選び出された類型特徴です。これについて、ここでは一々立ち入った説明は省きますが、いくつか例を挙げると、例えば、最初に挙げた「流音のタイプ」というのは、日本語ラ行子音の特徴で、日本語にはこのラ行子音が1種類しかなく、例えば英語などヨーロッパの諸言語に見られるr音とl音の区別というものがありません。この現象に着目すると、世界の言語はrとlを区別する「複式流音型」、それを区別しない「単式流音型」、さらにまたごく少数ですが、流音という音素を持たない「流音欠如型」という3つのタイプを立てることができます。

2番目の「形容詞のタイプ」、これは形容詞という語類が品詞として名詞に近い語類として位置づけられるか、それとも動詞に近い語類として振る舞うかという観点からの分類です。日本語のように形容詞が動詞に類した活用をし、そのまま述語として用いられるような言語が「形容詞用言型」、アルタイ諸語や古い印欧語のように、形容詞が名詞と同じような格変化をしたり、そのまま名詞として用いられるものが「形容詞体言型」として区別されます。

また4番目の「類別タイプ」、これは例えばドイツ語やロシア語などは、名詞に男性、女性、中性などのいわゆるジェンダーの区別があります。一方、日本語にはこのようなジェンダーという現象はないけれども、ものを数えるときに、人間ならば「ひとり、ふたり」、犬や猫なら「1匹、2匹」のように対象物の意味的カテゴリーに応じて違った「助数詞」というものを使います。このように、ジェンダーまたはクラスによって名詞自体を直接類別するタイプを「名詞類別型」、数詞などによって間接的に類別するタイプを「数詞類別型」として2つのタイプが立てられます。

また、「造語法的手段としての重複」というのは、日本語で名詞なら「山々」「国々」、形容詞なら「高々」「細々」、動詞なら「飛び飛び」「行く行く」などのように、語の一部ないし全体を繰り返す造語法で、これを多用するかしないかというのも言語の位置づけにとって重要な目印となるわけです。

このような類型特徴が世界言語の中でどのような分布を見せるか。ここではアフリカ・ユーラシア・オセアニアの諸言語に限って、調査されたそれらのデータを整理してみると、配付資料2ページ目の「表2」のような形となります。また、このようなデータに基づいて、それを世界言語地図上に概略的な分布図として表したいいくつかの例を、配付資料7ページ以下(図1から図3)に挙げてあります。

まず、「図1」をご覧ください。日本語のような単式流音型の言語は、ユーラシアではその大平洋沿岸部に集中的に分布し、しかもその分布圏は、ベーリング海峡を越えて遠くアメリカ大陸まで広がっています。これと全く同じような分布を示すのが、「図2」に示した形容詞用言型のタイプです。一方「図3」に示した類別タイプでは、日本語のような数詞類別型の言語は、その分布圏が1, 2の場合より狭まるけれども、やはりユーラシアからアメリカ大陸にかけて分布の広がりを見せています。

このような類型特徴の地理的分布に基づいて、日本語がどのように位置づけられるか、これをとりあえずユーラシア諸言語の範囲内で纏めてみると、最初の「表1」で示したような形になるわけです。

この表でご覧のように、ユーラシアの諸言語は「ユーラシア内陸言語圏」と「太平洋沿岸言語圏」という2つのグループに大きく分けることができます。

この中で「太平洋沿岸言語圏」は、地理的に「南方群」と「北方群」の2つにはっきりと分かれます。その南方群には、「ミャオ・ヤオ」、「タイ・カダイ」、「オーストロアジア」、「オーストロネシア」という4つの語族が含まれ、これらは全体として「オーストリック大語族」というような名称で纏めることができるでしょう。それに対して北方群というのは、日本海を取り囲む比較的狭い地域に分布する4つの言語です。すなわち、朝鮮半島の朝鮮語、日本列島の日本語とアイヌ語、列島北端の樺太と対岸のアムール下流域に残存するギリヤーク語がひとつグループとして纏まります。私はこれらの言語を「環日本海諸語」と名づけましたが、このように見てくると、日本語はまぎれもなく環日本海諸語の一員として位置づけられることが分かります。

一方、「ユーラシア内陸言語圏」は、それぞれ共有された一連の類型特徴に基づいて、その中心部に連続した広域分布を示す「中央群」とその周辺部で非連続な分布を示す「周辺・残存群」と名づけられるような言語群に分けられます。ただ、ここで注目されるのは、表のちょうど中間に置かれたチベット・ビルマ語族と漢語(つまり現在の中国語)です。この2つは「シナ・チベット語族」という名称でひとつの語族を形作ると見なされているわけですが、その中の漢語は、ここで扱った類型特徴から判断するかぎり、太平洋沿岸系の

言語とほとんど同じ特徴を共有し、従って、その中の一員と見なされるような性格を具えています。そのためこの言語群は、あたかも2つの言語圏の間で両者をまたぐような形になってしまうわけですが、漢語という言語がなぜこのような位置づけになるのかという問題が生じます。これについて、私の考えを結論的に申しますと、この言語はチベット・ビルマ系の言語と太平洋沿岸系の言語が、今からおそらく4千年ほど前に、黄河中流域のあたりで接触した結果生まれた一種の混合語（言語学の用語でいわゆる“クレオール”）と見なされる、というのが私の解釈です。これは古代の漢語自体の構造的な特徴、例えば、動詞や名詞に語形変化が全く欠けるとか、非常に変則的なSVO型語順を持つというような点からも、十分に証拠づけられるのです。

ところで、このような類型地理論から導かれた「太平洋沿岸言語圏」に関して、もうひとつ大事な点は、先ほどのいくつかの分布図を見ても分かるように、この言語圏がユーラシア大陸だけでなく、ベーリング海峡を越えて遠くアメリカ大陸まで広がっていることです。アメリカの先住民言語は、南北両大陸を通じて非常に複雑・多様な様相を示していますが、これらの類型特徴の分布をたどっていくと、この大陸の場合も、同じ特徴を共有する諸言語はこの大陸全体ではなく、どうやらその西側、つまり太平洋沿岸部に集中することが分かってきました。そこで、私は太平洋を挟んで2つの大陸の沿岸部に広がるこの大きな言語圏を「環太平洋言語圏」と名づけたわけです。

ただし、アメリカ大陸の場合、ここで扱ったような類型特徴だけでは、複雑に分布する諸言語のどのあたりに境界線があるのか、それを正確に見極めるのが非常に難しい、という方法論上の問題も出てきます。このような類型特徴というものは、それがどんなに根源的と見られるにしても、言語接触などによる伝播・拡散という現象を免れないからです。その意味で、ここで試みた類型地理論的アプローチは、やはり水深測量的な性格を脱し得ないと言えるかもしれません。

3 人称代名詞から導かれた世界言語の系統分類

このような限界を乗り越えるために、私が改めて取り組んだのが人称代名詞です。これは、「表1」の8番目に挙げた「1人称複数の包含・除外の区別」という現象に関連して、世界言語の人称代名詞のデータを集める中で次第に分かってきたことですが、この人称代名詞というものが、言語の遠い親族関係を探る上で、非常に有力な決め手となる。というのは、同じタイプの人称代名詞がしばしば語族の枠を超えて広域に分布することが明らかになったからです。

ユーラシアには、そのような広域に分布する人称代名詞として、ひとつは「太平洋沿岸型」、もうひとつは仮に「ユーロ・アルタイ型」と名づけられる2つのタイプが挙げられます。この中で太平洋沿岸型の人称代名詞は、その分布が先の「表1」の太平洋沿岸言語

圏と完全に一致します。さらにまた、この人称代名詞の分布をたどっていくと、アメリカ大陸の太平洋沿岸系言語の輪郭も、かなりはっきりした形で捉えることができます。ここでその地理的分布の全体像を纏めるみると、配付資料3ページの「表3」のような形で整理できます。

この表で見ると、太平洋沿岸型人称代名詞は、その現れ方がやや複雑で、1人称には基幹子音として k- を持つ形と n- をもつ形の2つがあって、言語によってそのどちらかが選ばれるか、あるいは同じ言語の中で2つが共存する場合は、k- 形は動詞に接して主語人称、名詞に接して所有人称、一方 n- 形は目的語人称を表すという形でその役割が分かります。身近には、アイヌ語の ku と en がまさしくそれに当たります。

また「包括人称」(これは従来の1人称複数の包含形に対して私の与えた新しい名称ですが)これにも2種類あって、基幹子音として唇音 w-/b- を持つタイプ(A型)と t-/d- を持つタイプ(B型)が区別されます。日本語の ware や朝鮮語の1人称複数の uri などは、w- を基幹とするA型の包括人称として位置づけられ、また日本語の1人称の are, ore などは、元の基幹子音 k- の消失形に由来すると見られます。また特に注目される点として、この言語圏の中の環日本海諸語では、いわゆる「待遇法」の影響によって、本来の2人称代名詞が全面的に失われました。その結果、日本語をはじめこれらの言語では、本来の人称代名詞の姿が大変見えにくくなっています。

一方、「ユーロ・アルタイ型」の人称代名詞は、その分布が先に見た表1のユーラシア内陸中央群とほぼ重なりますが、そこに含まれる言語群が若干異なります。この人称代名詞は1、2人称に関してはその現れ方が非常に単純かつ明瞭で、それぞれの基幹子音が1人称で m-、2人称で t- (またはそこから転化した s-) となります。その具体例については、配付資料4ページの「表6」をご覧ください。この人称代名詞の語族を超えた広域分布の様相は、表のデータを見ただけでも、一目瞭然と言ってよいでしょう。

このような人称代名詞の検証・分析によって、ユーラシアとアメリカだけでなく、さらに世界諸言語全体の系統分類を試みたわけですが、そこで到達したその結論の部分だけを整理してみると、配付資料の「表5」のような形で表すことができます。アフリカの古い土着系言語を除くと、世界言語の人称代名詞のタイプは、最終的にどうやら6つのタイプに収まるようです。ユーラシアには、この6つのタイプがすべて現れ、先の「表1」でユーラシア内陸部の「周辺・残存群」として纏められた諸言語は、1番から4番のいずれかのタイプに帰属します。また、この6つの人称代名詞の中から「太平洋沿岸型」と「出アフリカ古層A型」(= “アメリカ東部内陸系”)、そして「出アフリカ古層B型」(= “ナデネ・エスキモー系”)の3つがアメリカ大陸へと運ばれました。

ところで、この表にはその一番右の欄に、それぞれの言語群ごとに、それに対応すると見られた遺伝子の系統が挙げてあります。Y-DNA というのは Y 染色体遺伝子、mt-DNA はミトコンドリア DNA で、いろいろなアルファベット記号は「ハプログループ」と呼ば

れるそれぞれの遺伝子系統に振り当てられた名称です。日本語をめぐる言語系統論の議論の中に、Y 染色体やミトコンドリア遺伝子の名が出てくるのはいかにも唐突に感じられるかもしれませんが、実は、これが今日の私の話のもうひとつの大事なテーマです。

つまり、人称代名詞を含めて世界諸言語の類型地理論的考察から導かれた言語の遠い親族関係というものが、最近の分子生物学で「遺伝子系統地理論 phylogeography」と呼ばれる分野の研究成果とはたしてどのようなつながりを見せるのか、という問題です。ご承知のように、この分野はここ 10 年余りの間に急速な進展を見せましたが、出アフリカを果した現生人類が世界の各地へいつどのような形で移住したのか、こうした問題に従来とは全く違った新たな展望を切り開きました。

4 言語の系統と遺伝子の系統

言語の系統と遺伝子の系統との関連を探るためには、「ミトコンドリア」という女系遺伝子と「Y 染色体」という男系遺伝子の 2 つの道が考えられますが、ここでは主に男系の Y 染色体遺伝子の側からこの問題を取り上げてみたいと思います。ミトコンドリア DNA に較べて、Y 染色体の方がハプログループと呼ばれる遺伝子型の種類が少なく、それぞれの系統関係も比較的分かりやすいからです。

そこでまず最初に、Y 染色体遺伝子の系統樹として最近の専門家によって提示されているその概略図を見ておきましょう。配付資料 12 ページの「図 7」< Y 染色体遺伝子の系統略図 > をご覧下さい。

これで見るように、現生人類の Y 染色体遺伝子は、究極的に今から 8 万年(?)ほど前の単一の祖先(いわゆる Y 染色体のアダム)に遡ると言われます。現在、Y 染色体のハプロタイプと呼ばれる変異型は、アルファベット記号で A から T まで 20 個のグループに分類されていますが、この中で A と B の 2 系統を除く残りのすべては、出アフリカ系の M168 という祖形から分かれたと見られています。この根幹からまず、DE、C、F という 3 つの枝が分かれ、さらにこの F がもうひとつの根幹となって、ここから G、H、I、J、K などが分かれ、そしてこの K の分枝とされる M526 という幹から M、S、N、O、P というような枝が分かれたと見られています。

ここでは、特にユーラシアの大平洋沿岸部という地域に着目すると、ここで主要な役割を演じた Y 染色体遺伝子は、主に D、C、O という 3 つのハプログループに絞られます。この中で、D 系統は“YAP”と呼ばれる DE-M145 の一方の分枝で、アフリカの E に対して“アジアの YAP”とも呼ばれています。この D 系統と C 系統は、その発現年代がおよそ 5 万年前とされ、出アフリカ系では年代的に最も古い遺伝子と見られています。

これらの遺伝子系統は、それぞれいくつかのサブグループに分かれ、その中でも D と C は非常に特異な地理的分布を見せています。その概略を示したのが同じページの「図 8」

です。これで見ると、D 系統はその出現地域がインド洋のアンダマン諸島、チベットとその周辺地域、そして日本列島というように全くかけ離れた場所に散らばって、言語地理学の用語を借りれば、典型的な“周辺残存分布”の様相を呈しています(なお、D* のように アスタリスクの付くのは、特定のサブタイプに属さない「パラグループ」と呼ばれる枝を指します。アンダマン諸島でこの遺伝子は、「南アンダマン語」を話す「オンジェ」、
「ジャラワ」という小さな集団でほぼ 100% の出現率を見せています)。

一方、C 系統は D 系統より分布域が広いけれども、一部はメラネシアとオーストラリア、一部は東北アジア、そして残りは日本列島と南アジアというように、これもやはり周辺残存的な分布と言ってよいでしょう(この中で、日本列島の C1 と南アジアの C5 分枝はいずれも出現率が 5% 以下の希少なタイプです)。

ちなみに、ニューギニアとオーストラリアが合体した「サフルランド」と呼ばれる大陸にこの遺伝子を携えた集団が到来したのは、今からおよそ 4 万 5 千年前と見られています。一方、日本列島を含む東アジア地域への D、C 系統を携えた集団の到来は、おそらく 4 万年から 3 万 5 千年前頃と見てよいでしょう。これが東アジアに到来した最初の現生人類にほかなりません。日本列島に関して言えば、考古学的には関東の「立川ローム層」の最下層(第 10 層)に「先ナイフ型」と呼ばれる古い石器を残し、また長野県の野尻湖畔に「立ヶ鼻遺跡」と呼ばれるナウマン象などの屠殺址を残したのがこの集団と見られます。

ここで、D、C というこの古い遺伝子系統とつながるのは、はたしてどのような系統の言語だったのかという問題を考えて見ると、その最も有力な候補として、先の「表 5」に挙げた「出アフリカ古層 A 型」と名づけた人称代名詞が浮かびます。すなわち、アフリカではニジェル・コンゴ語族とアフロ・アジア語族、オセアニアではオーストラリアの原住民諸語とニューギニアの私の命名で「内陸高地語群」と呼ばれるグループ(いわゆる「サフル系」)が、いずれもこれらの遺伝子とつながっています。またチベット・ビルマ系言語を特徴づけるシナ・チベット型の人称代名詞は、出アフリカ古層 A 型と基幹子音の構成が全く同じで、ただその中で 1 人称と 2 人称が置き換わった形のシステムです。そのため「出アフリカ古層 B 型」と名づけられましたが、古層 A 型からの派生システムと見れば、これもまたこの古い遺伝子とつながる可能性が高いと言えるでしょう。

ところで、この古い遺伝子、特に D 系統は、後に見るように、日本列島にも非常に高い頻度で出現します。しかし、ここには出アフリカ古層系と見られる人称代名詞は見出されません。環日本海諸語とその外延に広がる太平洋沿岸言語圏と直接つながる遺伝子は、この C、D 系統より年代的にもっと新しい O-M175 と呼ばれる系統です。

この O 系統は、現在東アジアで最も優勢な Y 染色体遺伝子で、その中に沢山の枝を抱えています。主要なサブグループとして O1、O2、O3 という 3 つの枝が区別されます。そしてこれらのサブグループは、太平洋沿岸部の主要な言語群とそれぞれ密接なつながりを見せています。その概略を示したのが配付資料 12 ページ「図 8」の下の欄です。

これで見ると、太平洋沿岸系言語ととりわけ密接につながるのは、O2-P31 というサブグループです。しかもこの遺伝子は、O2a と O2b という2つの枝に分かれて、それぞれの分布圏が南と北にはっきりと分かります。そして O2a はこの言語圏の「南方群」と、O2b は「北方群」つまり環日本海諸語の分布圏とほぼ正確に一致します。

ただし、O2a によって特徴づけられる南方圏では、O 系統のもうひとつのサブグループ O1-M119 という枝がそこに加わって、南方群の中の特にオーストロ・タイ系の言語と密接なつながりを示しています。

この O1 と O2 に対して、O3 系統は、現状で見ると、中国大陸の漢語圏を中心に最も優勢な分布を見せる遺伝子ですが、O3 のサブグループ全体が漢語ないしシナ・チベット系というわけではありません。ただ、その中で O3e-M134 という分枝が、チベット・ビルマ系および漢語系集団と最も密接につながる遺伝子として位置づけられるのです。

なお、O3 と呼ばれる遺伝子は、その下に非常に沢山のサブタイプを派生させていて、それらの正確な分類や命名法はまだ確立されていません。図 8 に示した O3c, d, e というのも、遺伝子系統論の初期の名称をそのまま便宜的に使っています。

5 東アジア諸集団の Y 染色体遺伝子系統の分布

以上、Y 染色体遺伝子系統の全体像と東アジアで主要な役割を演じた D, C, O 系統についてのごく簡単な説明ですが、ここで東アジアの主な集団について、それぞれの遺伝子の構成がどのようになっているか、実際のデータに基づいて簡単に見ておきましょう。配付資料の 5 ページ〈東アジア諸集団の遺伝子的背景〉の「表 7」以下をご覧ください。

まず、「表 7」は東北アジアとシベリア諸集団のデータです^{*1}。これで見ると、モンゴル・ツングース系集団を特徴づける Y 染色体遺伝子は C3 系統で、そこにシベリア系の N 系統が加わります。ほかに、一部の集団で O 系統が出現していますが、その中の O2 はどうやら O2b で、わずかながら環日本海域とのつながりを窺わせます。一方 O3 系統は、明らかに漢語・漢民族圏からの流入で、漢語圏との地理的並びに社会的な“接近度”を端的に示しています。この中で中国領内に取り込まれた内モンゴルやヘジェン・オロチョンなどのツングース系集団は、現在ではほとんどその母語を失って、言語面での漢語化が進んでいます。

次に「表 8」は、主な漢語系集団の遺伝子データです^{*2}。この表で、O1, O2a, O3d は太平洋沿岸系、O3 の中のとりわけ O3e が漢語およびチベット・ビルマ系を特徴づける遺伝子です。O3 系統は、明らかに北方集団で出現頻度が高く、それに対して南方集団では、太平洋沿岸系のとりわけ O1 系統が長江下流域（浙江、上海など）の集団で高い出現率を

*1 出典：Xue et al. 2006, Jin et al. 2009.

*2 出典：Wen et al. 2004a.

示しています。

なお、北方集団の中では、甘肅の遺伝子頻度がやや特異ですが、これは元もとチュルク語を話していたイスラム系集団が言語的に漢語化した集団で、中国で回(ホイ)族と呼ばれています。ここに現れた O3 の出現頻度は、まさしくこのような言語置き換えの遺伝子面での反映と言ってよいでしょう。

次に、「表 9」のチベット・ビルマ系集団で注目されるのは、チベットから雲南地域にかけての一部の集団で、出アフリカ系の古い遺伝子 D 系統が飛び抜けて高い頻度で出現している点です*3。それに呼応して、O3 系統の出現率が集団によって極端に低くなっています。これがチベット・ビルマ系集団の遺伝子構成の最も特異な点ですが、この集団とその言語にとって最も本源的な遺伝子は、おそらく D 系統と見てよいでしょう。

それでは次に、太平洋沿岸言語圏に属する集団の Y 染色体遺伝子の出現状況を見てみましょう。まず「表 10」がその北方群つまり環日本海域です。この中の満州集団は、地域的にかつての環日本海域に属するけれども、現在その言語はもちろん環日本海系ではありません。かつて朝鮮半島北部から満州にかけて居住していた扶余・高句麗系の言語が環日本海系の言語だったと見られますが、現在は完全に消滅しました。またこの表には北方のアイヌやギリヤーク集団、またロシア領沿海州に分布するツングース系集団のデータが欠けているのが残念ですが、これらの地域の信頼に足る調査報告は、管見のかぎり、まだ出されていません。なお、この表で「日本 1, 2」などとあるのは、偏りを避けるために、それぞれ違った調査データを挙げたためです*4。

これで見ると、日本列島の遺伝子構成が最も特異な様相を示していることが分かります。その最大の原因は、その出現がほとんどここだけに局限される C1 と D2 の存在です。中でも D2 系統は、他のどのハプログループよりも高い出現率を示しています。一方、O 系統は、先のチベット・ビルマ系集団と同じように、D 系統の出現率と相反ないし相補的な関係となって、全体としての出現率が 60% 前後となっていますが、その中で O2b の出現率が最も高い、これが非常に大事なポイントです。

一方、朝鮮半島と満州の集団では、O 系統の出現率が全体で 80% 近くに達するけれども、その中で環日本海域を特徴づける O2b の出現率は 30% 前後で、O3 のそれよりも下回っています。O3 系統が 40% を超えるこのような高い出現率は、東アジアのシナ・チベット系以外の集団では、ほかにほとんど例を見ません。先の中国領内のツングース系やチュルク系イスラム集団で見たように、O3 系統の出現率が 30% かそれ以上に達した集団では、ほとんどの場合その言語が漢語によって置き換えられています。実際、この表の満州系集団は、現在ではその母語を完全に失って、事実上“漢族化”してしまいました。ところが、O3 系統の出現率が 40% を超えるこの朝鮮半島では、「朝鮮語」という紛れもな

*3 出典：Wen et al. 2004b, Cordeaux et al. 2004.

*4 出典：Nonaka et al. 2007, Hammer et al. 2006, Kim et al. 2011, Katoh et al. 2005.

く環日本海系の言語が今なおしっかりと維持されています。これはほとんど奇跡に近い、極めて珍しいケースと言ってよいでしょう。しかも、朝鮮の O3 系統の中で O3e の出現率は、朝鮮 1 の調査で 27% にも達し、先に見た漢語系集団とほとんど変わりません。朝鮮半島と漢語系集団との違いは、ひとえに O2b の出現にかかっているのです。

環日本海域の O 系統についてもうひとつ付け加えると、ここでは南方群を特徴づける O1 および O2a がほとんど流入していません。これもまた、日本語の系統を考える上で、非常に大事なポイントです。

なお、この表に現れる C3 系統と N 系統は、いわゆる北方系の遺伝子です。この 2 つは朝鮮半島と満州ではかなりの出現率を示していますが、日本列島ではほとんど見るべき役割を演じていません。長年、日本の学会で有力視されてきた日本語のいわゆる“ウラル・アルタイないしアルタイ語起源”というような学説は、遺伝子系統論の側からは全く支持されないといいましょう。

次は、「表 11」から「表 13」まで、太平洋沿岸南方群の主な集団の遺伝子データです*5。この中で、ミャオ・ヤオおよびオーストロアジア系集団では、その原郷地を遠く離れた集団ほど、遺伝子の多様性が失われて、集団の中核を担う O2a の出現率が高くなるのが注目されます。ちなみに、インド東部のムンダ系集団で、O2a 以外の遺伝子はすべてインドの土着系(特に H 系統)で占められています。

続く「表 12」と「表 13」は、オーストロ・タイ系のデータですが、ここでは O2a に代わって O1 系統が明らかに優勢となっています。また「表 13」で注目されるのは、インドネシアの西と東で Y 染色体遺伝子の出現が極端な相違を見せる点です。その境界はバリ島とフローレス諸島の間であって、それより東の地域では、東アジア系の遺伝子つまり O 系統の出現率が急減し、代わってメラネシア系の遺伝子の出現率が著しく高くなっています。表の右側に配した C2, M, S がそのメラネシア系です。この地域(つまりオセアニア)では、オーストロネシア系言語の担い手が Y 染色体つまり男系遺伝子から女系のミトコンドリア DNA へと一方的に替ってしまったのです(松本 2010:693ff.)。

最後に、「表 14」に挙げたデータは、長江流域の古人骨から検出された Y 染色体遺伝子についての最近の貴重な研究成果です(Li et al. 2007)。ここで「歴史時代」とあるのは、概ね漢代に属する古人骨です。これで見ると、長江下流域の人骨で検出に成功した事例は、先史時代から漢代に至るまですべてが O1 系統で、これらの文化を担った集団がオーストロ・タイ系に属していたことをはっきりと示しています。一方、長江中流域から検出された遺伝子は、そのほとんどが O2a と O3d に属し、これは紛れもなくオーストロ・ミャオ系の遺伝子です。いずれも、長江流域に発祥した稲作文化の担い手集団でした。

この報告には、他に黄河中流域の龍山文化に属する人骨のデータも含まれていて、それ

*5 出典 : Kumar et al. 2007, Chaubey et al. 2011, Li et al. 2011, Delfin et al. 2011, Karafet et al. 2010.

を見ると、ここには太平洋沿岸系の遺伝子が全く現れず、後の漢語系集団を特徴づける O3 系統で占められています。龍山文化は後の夏王朝や殷・周王朝の母胎となった文化です。北方の黄河流域と南方の長江流域との間で、その遺伝子の構成が全く異なることがこの表からはっきり読みとることができます。

なお、参考までに、太平洋沿岸南方群に属する諸言語および各集団の推定された古い原郷地とこれらの言語の現在の分布図を、配付資料 11 ページ (図 5 と図 6) に挙げておきました (詳しくは松本 2007:277ff. 参照) 。

以上、東アジア諸集団の Y 染色体遺伝子の出現状況についての概要です。これを念頭に置いた上で、これらの遺伝子系統の現在の地理的分布がどのような形になっているか、この分野の専門家による分かりやすい分布図が最近公表されているので、ここで関連する部分を若干の修正を加えて、配付資料の最後の 2 ページに挙げてあります。

まず「図 9」をご覧ください。これは Y 染色体の主要なハプログループ全体の分布図です。図の上段 A, B, E がアフリカ、2 段目の D, C, O がアジア・太平洋地域、3 段目の N, Q, R が主にユーラシア内陸部 (およびシベリア) に広がったグループです。今日の配付資料の冒頭で示した類型地理論的アプローチによる「太平洋沿岸言語圏」と「ユーラシア内陸言語圏」という言語上の 2 大区分は、これを遺伝子の側から眺めると、概略的に、それぞれ O 系統と R 系統の分布圏に対応していることが分かります。

次の「図 10」は、3 つのサブグループを含めた O-M175 の分布図です。これを見ると、中央の O2 系統の分布域が太平洋沿岸言語圏のそれとほぼ正確に一致することが分かります。しかもこのグループは、南方の O2a の分布域と北方の O2b との間でその境界がはっきりと分かれ、双方の間で交流・混合の現象がほとんど見られません。これは O2a と O2b を携えた 2 つの集団、そしてまたこの遺伝子によって特徴づけられる南・北 2 つの言語圏の分岐した年代が非常に古いことを示唆しています。

一方、O1 系統は、前にも述べましたが、O2a を基調とする南方群の中の有力な支脈であるオーストロ・タイ系集団に特化したサブタイプです。大陸の沿岸部ではタイ・カダイ系、また特にその下位枝 (O1a2-M50) が台湾から東南アジア島嶼部に広がったオーストロネシア系集団を特徴づけています。「タイ・カダイ」と「オーストロネシア」という 2 つの語族の緊密なつながりを示す遺伝子型と言ってよいでしょう。

ちなみに、この図で O2b の出現率が最大で 30% となっているのは、現状で O2b の分布の中心と見られる日本列島で、列島に温存された古い遺伝子 D 系統によって、O2b の出現率が相対的に低められているという事情が関係しています。

ところで、これらの分布図を全体として眺めると、太平洋沿岸言語圏を特徴づける O2 および O1 の分布圏は、その大陸部を中心に O3 系統 (とりわけ O3e-M134) によってその表面を大きく覆い隠されていることが分かります。これは、言うまでもなく、過去 2 千

年近くにわたって行われた漢語圏の拡散・膨張の結果にほかなりません。現在、中国本土の中心部では、古い太平洋沿岸系の言語はこの新興言語に呑み込まれて完全に姿を消してしまいました。

6 太平洋沿岸系集団の環日本海域への到来時期

最後に、O2b（あるいはむしろその祖型となった O2-P31）という遺伝子を携えた集団がいつ頃この日本海域に到来したか、という問題を取り上げてみましょう。

戦後日本の人類学や考古学界では一時期、日本人の成り立ちに関していわゆる“二重構造説”とされる学説が流行しました。これによると、日本の縄文時代と弥生時代の間に集団的そしてまた言語的にも大きな転換があって、現在の日本人は弥生時代の開幕期に外部から稲作や金属器使用など新しい文化を携えた渡来系集団にその直接のルーツを持つとされてきました。日本人の祖先が稲作を携えて南方から海を渡ってやってきたという考えは、柳田国男氏の『海上の道』という著書（1961）に遡る根強い学説ですが、これを現在の遺伝子系統論の側から見るとどうでしょうか。

まず第一に、長江流域に発祥した稲作民（つまり太平洋沿岸南方系集団）の遺伝子（O1, O2a）は、先に見たように、日本列島にはほとんど流入していません。環日本海域を特徴づける O2b は、中国大陸を含めた南方の稲作圏とは全くつながらない遺伝子です（これはもちろん稲作の伝来が問題となるような次元での話ですが）。また日本に新しく流入した可能性の高い O3 グループも、地域的には黄河流域に発祥する中国北方文化圏を特徴づける遺伝子です。

弥生時代以降に外部から日本に流入した Y 染色体遺伝子としては、O3（特に O3e-M134）が最も有力な候補と考えられますが、この遺伝子の流入は、弥生時代よりもむしろ古墳時代に入った紀元 3～4 世紀以降ではなかろうか、と私は見ています。日本ではこの時期から、朝鮮半島を通じて漢語・漢文化圏との接触・交流が急速に強まってきます。『日本書紀』などで「イマキのアヤヒト」とか「カワチのフミヒト」などと呼ばれたいわゆる帰化人や百濟滅亡に伴う大量の難民などが、そのような遺伝子流入に大きな役割を演じたに違いありません。

O2 系統によって特徴づけられる太平洋沿岸系集団の環日本海域への到来の時期については、アメリカ大陸への人類移住が重要な鍵を握っています。すでに述べたように、太平洋沿岸系の言語はアメリカ大陸へも運ばれ、南北両大陸の大太平洋沿岸部に大きな分布圏を作っています。アメリカ大陸への移住を成し遂げた人類集団の少なくとも一部は、太平洋沿岸言語圏の北方域つまり環日本海域をその出発地としたと見なければなりません。

アメリカ大陸への人類移住の時期やルートに関して、最近の研究によると、シベリア東部からベーリンジアを歩いていわゆる“直行便形式”で行われたのではなくて、途中にか

なりの滞留期間があった。つまり、最終氷期最寒冷期以前にシベリアに進出していた人類集団が、最寒冷期の到来によってどうやらベーリンジアに閉じこめられたらしい。ここが極地圏で残された居住可能なほとんど唯一の避難地だったからです（配付資料の「図4」参照）。そしてこの地で3千年ほどの時期を過ごした後、温暖化の到来を待って、一部は太平洋の沿岸ルート、一部はやや遅れて内陸の“無氷回廊”を伝ってアメリカへの移住を成し遂げた、というシナリオが有力視されています。

ちなみに、アメリカに運ばれた遺伝子が、ミトコンドリアとY染色体の双方で、旧大陸のそれに較べてその構成が非常に単純化され、また特にその男系遺伝子が極端に偏った形になったのは、閉ざされたベーリンジアでの長期にわたる滞留中に、いわゆる“ボトルネック効果”によって遺伝子の多様性が急激に減少したためと見られます。ここで沿岸系言語の担い手も、どうやらオセアニアと同じように、男系遺伝子から女系のミトコンドリアDNAへと一方的に転換してしまったようです（松本 2010:692ff.）。

いずれにせよ、上に述べたアメリカ移住に関するシナリオが受け入れられるならば、太平洋沿岸系言語を携えた集団が日本列島を含めた環日本海域へ到来した時期は、当然、最終氷期最寒冷期以前と見なければなりません。とすれば、遅くとも今から2万5千年前あたりになるのでしょうか。Y染色体のD系統を携えた最初の人類集団の日本列島への到来を4万年～3万5千年前とすれば、O2系統を携えたこの第二次集団の到来は、それより1万年ほど遅れたと見てよいでしょう。考古学的には、「石刃技法」と呼ばれる新しいタイプの石器を列島にもたらしたのがこの集団とつながるかもしれません。

このように見てくると、日本語および日本人のルーツは、縄文時代をはるかに越えて、少なくとも2万年以上前まで遡るという結論に導かれます。従ってまた、仮に日本人の二重構造説というような見方が成り立つとすれば、それは縄文と弥生を隔てる2～3千年前というような間近な時期ではなくて、Y染色体のO系統を携えた太平洋沿岸民が列島に到来した2～3万年前あたりまで遡らなければなりません。

この新しい集団の到来によって、D系統と共に日本列島にもたらされた出アフリカ古層系と見られる言語は、新来者の言語すなわち太平洋沿岸系言語によって完全に置き換えられました。けれども、“アジアのYAP”と呼ばれるこの古い遺伝子自体は、新来の遺伝子によって駆逐されることもなく、そのまま温存されたのです。アフリカに直結するこの古い遺伝子が4万年もの長きにわたってこの列島内で存続し、しかも現在の日本人の男系遺伝子の中でその出現率が一番高い（関東地方の男性ではその出現率が48%という調査報告もあります）、これもまた奇跡的と言ってよいでしょう。いずれにしても、日本語のルーツはその背後に2万年以上に及ぶ悠遠な過去を潜ませていることを、ここで改めて強調しておきたいと思います。

以上で、本日の講演を終わらせて頂きます。ご清聴ありがとうございました。

【引用文献】

- 松本克己 2007 『世界言語のなかの日本語：日本語系統論の新たな地平』東京：三省堂
- 松本克己 2010 『世界言語の人称代名詞とその系譜：人類言語史5万年の足跡』東京：三省堂
- Chaubey, G. et al. 2011 Population genetic structure in Indian Austroasiatic speakers: the role of landscape barriers and sex-specific admixture. *Mol. Biol. Evol* 28(2): 1013–1024.
- Chiaroni, J. et al. 2009 Y chromosome diversity, human expansion, drift, and cultural evolution. *PNAS* vol.106, no.48:20174–20179.
- Cordeaux, R. et al. 2004 The Northeast Indian passageway: a barrier or corridor for human migrations? *Mol. Biol. Evol.* 21(8):1525–1533.
- Delfin, F. et al. 2011 The Y-chromosome landscape of the Philippines: extensive heterogeneity and varying genetic affinities of Negrito and non-Negrito groups. *European Journal of Human Genetics* 19:224–230.
- Hammer, M.F. et al. 2005 Dual origins of the Japanese: common ground for hunter-gatherer and farmer Y chromosomes. *Journal of Human Genetics* 51:47–51.
- Jin, H.J. et al. 2009 The peopling of Korea revealed by analyses of mitochondrial DNA and Y-chromosomal markers. *PlosOne* vol.4(1), e4210.
- Karafet, T.M. et al. 2010 Major east-west division underlies Y chromosome stratification across Indonesia. *Mol. Biol. Evol.* 27(8):1833–1844.
- Katoh, T. et al. 2005 Genetic features of Mongolian ethnic groups revealed by Y-chromosomal analysis. *Genetics* 346:63–70.
- Kumar, V. et al. 2007 Y-chromosome evidence suggests a common paternal heritage of Austro-Asiatic populations. *BMC Evolutionary Biology* 7:47.
- Li, H. et al. 2007 Y chromosomes of prehistoric people along the Yangtze River. *Human Genetics* 122:383–388.
- Li, H. et al. 2008 Paternal genetic affinity between western Austronesians and Daic populations, *BMC Evolutionary Biology* 8:146.
- Nonaka, I. et al. 2007 Y-chromosomal binary haplogroups in the Japanese population and their relationship to 16 Y-STR polymorphisms. *Annals of Human Genetics* 71:480–95.
- Wen, B. et al. 2004a Genetic evidence supports demic diffusion of Han culture. *Nature* 431:302–305.
- Wen, B. et al. 2004b Analyses of genetic structure of Tibeto-Burman populations reveals sex-biased admixture in Southern Tibeto-Burmans. *American Journal of Human Genetics* 74: 856–865.
- Xue, Y. et al. 2006 Male demography in East Asia: A North-South contrast in human population expansion times. *Genetics* 172:2451–2459.