

公開講演会要旨

日本語の起源を探る

産能大学* 安 本 美 典

(1995年11月1日, 統計数理研究所 講堂)

1. 「古極東アジア語」の設定

最近『ラングウィッジ (Language)』誌など, 外国の言語学の雑誌には, 言語と言語の間の近さの度合いを, 数字ではかる研究が相次いで発表されている. 確率論や統計学, 因子分析法などを基礎とするものである.

2つの言語を確率論や統計学を用いて比較する方法では, 2つの言語を比較した場合の実際の統計的一致率が, 偶然による一致の率を超えているかどうかを計算する. 1つ1つの単語を比較するだけでは, 2言語間の一致が偶然のものであるか否かは判別しにくく, 統計的に扱うことによってのみ, 偶然以上の一致が浮かび上がる.

私たちも, このような研究に刺激され, 日本語と60ほどの世界の言語との関係の度合いを数字ではかってみた.

調査したのは「数詞」, 手・口・鼻などの「身体語」, 鳥・雲・水など, どのような民族集団でも, 必ずそれに当たる単語をもっているような項目からなる「基礎百語」および「基礎二百語」である. このような「基礎語彙」は, どの言語でも, 時間の経過に対する抵抗力が強い. すなわち, 古い時代の語が, それほど変化せずに残る傾向が強い.

表1. 日本語と朝鮮語との身体語.

	日 本 語		朝 鮮 語	
	現 代 日 本 語	上 古 日 本 語	現 代 朝 鮮 語	中 期 朝 鮮 語
手	te	tē	son	sōn
足	'asi	asi	par	pār
鼻	hana	řana	k'o	kō(h)
目	me	mē	nun	nūn
口	kutʃi	kuti	'ip	'ip
歯	ha	řa	'i	nī
耳	mimi	mimi	kwi	kūi
毛	ke	kē	t'or	t'ər
頭	'atama	kasira	mori	mərī
舌	sita	sita	hyo	hyō
腹	hara	řara	pe	pēi
背	senaka	se	tuŋ	tuŋ

* 経営情報学部: 〒259-11 伊勢原市上粕屋1573.

表1は、現代日本語と上古日本語（奈良時代の日本語）および現代朝鮮語の「身体語」を示したものである。現代日本語と上古日本語とは、1000年以上のへだたりがあるにもかかわらず、それほどの変化をみせていないことがわかる。また、日本語と朝鮮語の身体語がかなり異なっているのもわかる。

「基礎語彙」は、言語の核心部分である。だから、言語を手がかりとすれば、数千年の時間の霧の中に晴れ間を見いだし、文献以前の民族の歴史をかいま見ることができる。

では、こうした新しい日本語の起源探究方法によるとき、日本語と偶然とはいえない関係を示す言語は存在するのであろうか。数字ではかってみると表2に示すように、日本語と偶然といえない関係をもつ言語がいくつか浮かび上がる(表2)。

まず、日本語、朝鮮語、アイヌ語の3つの言語は、「基礎百語」「基礎二百語」の調査で、相互に確率論的に偶然とはいえない関係を示し、ひとつのまとまりを見せている。朝鮮語は日本語とそれほど近くはないが、確率論的には偶然以上の関係が見いだせる。私たちは、これらの言語を「古極東アジア語」系の言語と名づけた。

古極東アジア語系の言語では、語彙の近さでまとまりを見せるだけではない。①語頭に2つ以上の子音がこない ②「r」と「l」の音の区別がない ③本来は清音と濁音の区別がなかったらしい ④二重母音を避ける傾向がある ⑤語の平均の長さはほぼ2音節である ⑥基本的な語順が一致する ⑦日本語の「てにをは」に当たるものをもつ、など音韻上、文法上でも共通の性質をみせる。

2. 身体、植物関係語は江南から

「古極東アジア語」は、近隣のアルタイ諸言語の祖語的なものから、極めて古く分離したもの

表2. 上古日本語と偶然以上の一致のみられるもの（基礎200語）。

言 語	一致が偶然によってえられる確率	一致数	註
東 京 方 言	0.000000 ^{.....}	157	一致が偶然によってえられる確率は、ほぼゼロである(0.0001%水準で有意)。
首 里 方 言	0.000000 ^{.....}	121	
インドネシア語	0.000145 ^{.....}	57	0.05%水準で有意。
カンボジア語	0.000544 ^{.....}	57	0.1%水準で有意。
中期朝鮮語	0.000663 ^{.....}	53	
ロ ロ 語	0.005649 ^{.....}	53	普通の統計学の基準では、1%水準で有意である。
中国語北京方言	0.008965 ^{.....}	50	
ネ パ ー ル 語	0.010315 ^{.....}	51	普通の統計学の基準では、5%水準で有意である。
ベ ト ナ ム 語	0.010467 ^{.....}	53	
モ ソ 語	0.024295 ^{.....}	52	
シンハリーズ語	0.031032 ^{.....}	47	
ベンガーリー語	0.037077 ^{.....}	49	
ビ ル マ 語	0.047021 ^{.....}	45	
タ ヒ チ 語	0.048605 ^{.....}	37	

表3. 日本語とビルマ系諸言語との身体語の一致.

	日 本 語		ボ ド 語 群				ナガ語群	クキ・チン語群	ビルマ語群
	上古語	首里方言	Bodo	Mech	Lälung	Dimāsā	Taungk'ul	TaungΘa	Tavoyan
手	tē	tii	ākhai	nākhai	$\begin{cases} iyā; \\ ja \end{cases}$	yao	pān	kut	lā ^k
足	asi	hwisja	$\begin{cases} āfa; \\ atheng \end{cases}$	nātheng	ja-thang	yēga	△ p'ei	hā-pā	ke pwa
鼻	fana	hana	ganthang	guntung	gung	gōng	nā-taŋ	○△ h ^a ra	○△ hnā kōi
目	mē	mii	○△ mēgan	○△ mōgan	○△ mu	○△ mū	○△ mik	○△ mi	○△ myi'si
口	kuti	kuci	○△ khūgā	○△ khuga	○△ khu	○△ kū	○△ k'a-mor	○△ kā	p ^a zā ^t
齒	fa	haa	○△ hāthai	○△ hāthai	○△ hā	○△ hātai	○△ hā	○△ hā	θwā
耳	mimi	mimi	khāmā	khumā	khānjur	kamao	k'a-nā	nā	nāpwā ^t
毛	kē	kii	○△ khenai	○△ khēnai	○△ khu-ni	○△ $\begin{cases} kamai; \\ kanai \end{cases}$	○△ kui-sam	tsam	sān bān
頭	kasira	çiburu	○ khārā	○khor	○khāpāl	○ kōrō	○kui	lu	○ kōh
舌	sita	šiba	○△ sila	○△ sālai	○△ si-li	○△ shalai	ma-le	le	○△ šā
腹	Hara	'wata	udoi	udui	○pu-mā	○ hō	△wuk	△ wān	△wun
背	sō;se	nagani	bikhung	bikhungā	lāngal-pathi	○ shīma	k'mu-k'or	kuñ-ru	△ nō'ja
上古語との語頭音の一致数			7	6	8	8	5	4	4
首里方言との語頭音の一致数			5	5	6	5	6	5	5

○印は上古語との一致を, △印は, 首里方言との一致を示している. なお, グリアンスンの表記では, jaはジャの音をあらわす.

のようで, 現在のアルタイ諸語との間にはかなり大きな溝がある.

日本語の形成にあたっては「古極東アジア語」が核をなしているらしいということがわかった. しかし, 身体語, 数詞などは朝鮮語やアイヌ語ではうまく説明できない.

そもそも, 「やま」「かわ」「とり」など普通の日本語の基礎語は2音節の語が多いのであるが, 「め」「は」「て」「け」など身体語には1音節のものが多い. 「みみ」「ほほ」「もも」などは, 「おめめ」「おてて」などというのと同じく, 本来は1音節語であったものが, 1音節では日本語として発音しにくいと, 音を重ねて2音節語にしたのではないかと疑われる形をしている.

ところで, ほとんどの単語が1音節からなる, これらの語は, 「やま」「かわ」といった普通の日本語の基礎語彙の上に, ちょうど油が水に浮くように浮かんでいる.

ビルマ語系の諸言語の中には, 身体語や植物関係の語などにおいて, 日本語と明らかに偶然とはいえない関係を示すものがある. これらの語は, 弥生時代のはじめごろ, どうやら稲作とともに, 当時中国の江南地方にいたビルマ系の言語を用いる人々によってもたらされたものらしい(表3).

3. 系統論のモデルは成立しない

以上のような比較のために用いた方法やデータについて, ここでは詳しく触れるだけの余裕がない. 『日本語の成立』(講談社現代新書), 『日本語の誕生』(大修館書店), 『日本語の起源を探る』(PHP研究所)といった3冊の拙著には, やや詳しく記してあるので, 興味のある方は参照していただきたい.

以下では, その結論のみをまとめる.

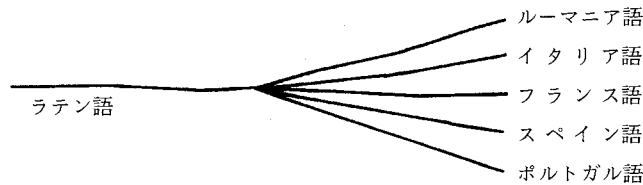


図1.「分裂論」のモデル.

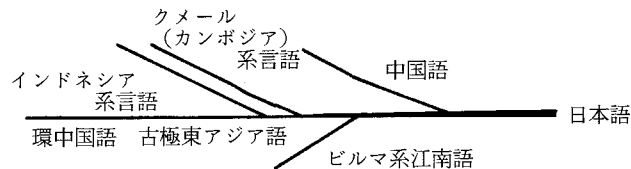


図2.「流入論」のモデル.

たとえば現在のスペイン語、フランス語、イタリア語、ポルトガル語、ルーマニア語などは、いずれも古代ローマで使われたラテン語から分裂し、独自に発達してきたものである(図1). この場合のラテン語のように起源になった言語を「祖語」という。そして同一の祖語から分かれた言語を「同系語」、あるいは「2つの言語は系統が同じ」という。

このような「系統論」は、1つの祖語から多くの言語が分かれて出ること想定している。すなわち、1つの源から出た流れが、多くの支流に分かれていくモデルである。印欧語のように、歴史上ほとんど常に他の言語に対して文化的に優位性をもち、ある源からあふれ出る形で発展してきた言語について、最もよく成立するモデルといえる。

しかし、日本語は全く逆に、多くの言語が流れ込む形で成立したと見なさなければならない。これは、多くの川が注ぎ込んで大河となるように、水源は1つとは限らないとするモデルである。このモデルのほうが、日本の地理的位置から考えても、歴史時代に入ってからわが国に多くの異質の文化が流れ込んだことを考えても、自然であると思える。

自然、という以上に、このようなモデルを考えないことには、日本語がいくつかの異質の言語と、統計的に偶然以上の一致を示す理由が説明できない。日本語と偶然以上の一致を示す言語をもとにつくった日本語成立についてのモデルが図2である。

結局、「日本語の起源」において成り立つのは、「系統論」ではなく、どのような言語が流れ込んだかの「流入論」、または、いかにして日本語が成立したかの「成立論」であると思われる。すなわち、日本語の成立に当たって流れこんだ言語としてどのようなものが考えられるか、そしてそれらの言語が日本語の成立にどう寄与したか、ということである。

印欧語は、ある源からあふれ出る形で発展したが、日本語は多くの言語が注ぎ込む形で成立している。この違いが、異なる形のモデルを要求するのである。

4. アイヌ語は古代の特徴残す?

日本語の形成について、現在の私の考えは次のようなものである。

(1) 今から5000～1万年前ごろ、ユーラシア大陸の内部から、東北部に押し出された言語があった。その言語を古極東アジア語と仮称するが、それは、アルタイ系諸言語の祖語と近縁関係にあったとみられる。

(2) 古極東アジア語を用いる人々は、朝鮮、北海道、そして北海道以外の日本に分かれて住み、それぞれの言語を独自に発展させて、朝鮮語、アイヌ語、日本語になっていった。3つの言語に共通の「語順」などは、古極東アジア語に由来するものとみられる。

(3) 朝鮮語はその後アルタイ諸言語からの影響を受け続け、アイヌ語は古極東アジア語の古代的特徴を比較的多く残したかとみられる。日本語は、その後も朝鮮語の影響を受け続ける一方で、縄文期にはインドネシア系言語、カンボジア（クメール）系の言語の流入を受け、それらが一般的な基礎語彙をもたらした。

(4) 紀元前2、3世紀ごろ、弥生時代の始まる前後に、おそらく稲作などとともに、江南からビルマ系の言語が日本に押し寄せてきたと考えられる。身体語、数詞、代名詞、植物関係の語が、これらの言語の流入によってもたらされた。

(5) 紀元前後から約2000年にわたり、中国語が流入し、多数の文化的語彙をもたらし、現在の日本語が形成されていった。