

『計量国語学』アーカイブ

ID	KK300201
種別	論文
タイトル	外行語世界分布の国別因子分析 —Google Trendsによる傾向—
Title	Country-wise Factor Analysis of Worldwide Distribution of “Lent Words”: The Tendencies by Google Trends
著者	井上 史雄, 柳村 裕
Author	INOUE Fumio, YANAGIMURA Yu
掲載号	30巻2号
発行日	2015年9月25日
開始ページ	73
終了ページ	97
著作権者	計量国語学会

論文

外行語世界分布の国別因子分析

—Google Trends による傾向—

井上 史雄 (国立国語研究所)

柳村 裕 (国立国語研究所)

要旨

外行語, つまり海外に進出した日本語について, インターネットの検索エンジン Google Trends を使って, 世界地図を作ることができる. その数値データはダウンロードできるので, 整理した上で因子分析にかけて, 多数地図の大傾向を把握した. 外行語を多く使う地域としてはアメリカ, 西欧, 東南アジアが際立つが, 国家のグループとしては中南米が他の国々と異なった傾向を見せた. Google Trends のデータによって, 従来不可能だった外行語の世界分布が解明できた.

キーワード: 外行語, 外来語, 因子分析, ゲーグルトレンド, 地理言語学

1. 外行語の研究史

外行語, つまり海外に進出した日本語について, インターネットの検索エンジンを使って, これまで見たことのない世界地図を作ることができる. これまで数百枚の世界言語地図を出力した. 個々の単語により分布パターンは違うが, 多数の地図を通してみると, 一般的な傾向も見られた. Google Trends の数値データはダウンロードできるので, 整理した上で因子分析にかけると, 多数地図の大傾向を把握できる. 実行したところ, 世界地図からの視察による予測とほぼ一致したが, 一部違った結果が得られた. 外行語を多く使う地域としてはアメリカ, 西欧, 東南アジアが際立つが, 国家のグループとしては中南米が他の国々と異なった傾向を見せた. 単語としては, 伝統的な日本文化を表わす単語と, 最近の若者向けのポップカルチャーを反映する単語が分離された. 南米は伝統的な日本文化を表わす単語に結びつき, 西欧は若者向けのポップカルチャーを反映する単語と結びつく. Google Trends のビッグデータによって, 従来不可能だった外行語の世界分布が解明できた. 本来の意味の地理言語学 *geolinguistics* の研究がこれによって可能になる.

「外行語」は外来語の対として名付けられたもので (三輪 1977), 借用語の一類型として日本語研究にとっては便利な名称なので, ここで用いる. 個別の報告はあるが, 全世界を見渡す研究は, 最近まで不可能だった. 多様なアプローチがあり, 電子辞書データが普及したので, 受け入れ国 (言語) ごとにはいくつかの言語で比較が可能だが (Canon 1996, 加藤・熊倉 1999, 早川 2003, 2006, 原口他 1998, 彭 2013), 調査の規模や時期の違いがあり, 単純に比較することはできない. またこれまでの研究は個別の語を扱うものが多く, 多数の語を計量的に扱う試みは多くない. ましてや全世界を扱った研究は限られ

ている(加藤他 1999)。近年英語と中国語を見渡した論考が出た(小林 2015)。術語としては「外行語」が早い¹(三輪 1977)、書名としては各種ある。ただし多くは「日本語」もしくは“Japanese”を含むもので、理論的考察の範囲を命名段階で限定する。後述のように言語間の借用関係は全世界の諸言語の相互関係としてとらえられるべきである。従来の研究は、研究手段の制限から某言語に入った(一言語または複数言語から借用された)単語が主だった。現在は様々な電子データが普及したおかげで某言語から他の複数言語に広がった単語の研究も可能になった。今回は日本語からの借用を扱い、日本語で書いたので「外行語」という術語を用いたが、言語的限定を含まない術語がいい。さらに単語だけでなく、音韻や語構成・形態論 morphology、構文論 syntax の相互影響も扱うことが望ましい。広い範囲を含みうる術語として“Borrowing”が候補になるが、日本語の術語「借用」と同じく方向性をはらむ。その逆方向の「貸付」を際立たせるためには、“Lending”が考えられるが、その論考はまだ少ないので、術語として確立していない。

個々の言語現象の世界分布は、南米起源の作物名や「茶」の発音の地図化などの形で、個別の発表があった。また言語学者も類型論の隆盛に従って、発音や文法に関わる現象を世界地図に表わしはじめた(WALS: World Atlas of Language Structures Online)。一方世界約 30 カ国の外行語 4 語の使用状況が日本語観国際センサスによって明らかになった(新プロ「日本語」総括班 1999, 国立国語研究所 2002)。

これらに加え、世界各国での語の使用状況がインターネット検索エンジンにより分かるようになった。中でも Google Trends では、思いついた単語について瞬時に地図化できる。しかも数表をダウンロードして集計できる。これを利用して、多数の単語、多くの国家について集計した研究もいくつかある(井上 2011a, 2012b, 2012c, 2013a, 2013b, 2013c, Inoue 2012a, 2013d)。ただしこれまでの研究は数表の行列の合計値を利用するのがせいぜいで、全体を多変量解析にかけるものではなかった。本論文では多変量解析の段階について論ずる。これまでの諸研究のまとめに相当するものである。いわゆる文化交流が借用に影響することが分かっただけでなく、経済交流が深く関わることが読み取れた。経済言語学の新たな領域として位置付けることも可能である(井上 2011b)。また本稿末尾で述べるように、将来のグローバルな歴史地理学的言語研究への序説となりうる。

2. Google データの研究手法

データの収集には検索エンジン Google Trends を用いた。空間的な地図と時間的なトレンドグラフとして出力されるが、手順を踏むと数表としてもダウンロードできる²。ただ

1 井上(2013c)で従来の用語を整理して提示した。そのほかに、小林(2015)によれば、これまで次のような用語が使われている。“Japanese borrowings”(Cannon 1981)、“Japanese loanwords”(エバンス 1990)、「日語借詞」(孫 1957)、「日語来源的外来詞」(高・劉 1958)、「日本語借用語」(沈 2008)。他に次のものもある:“JOL”(Japanese Origin Loanword)(ロング 2015, ロング他 2015)。なお井上(2013a)の注では「輸出語」「輸入語」が便利だと書いたが、“exported word”は術語としてはふさわしくない。対としての“imported word”が確立していないことが一つ、語の輸出先に何の見返りもないことが商品輸出と違うことが一つである。ただし 2 番目の理由は強く主張できない。なぜなら「借用 borrowing」と呼ぶ場合においても、借りた側が返すことがほとんど期待できないからである。

2 井上(2012b)および井上・柳村(2015)に手順を述べた。“グーグルトレンド”または“Google Trends”で検索して、ホームページにたどりつき、世界地図をクリックすると検索窓が出る。そこに任意の語を入れると、最近数年間の使用状況と世界地図が出る。地域を Worldwide から変えれば国ごとの地域

し数値は扱った語の最高使用率の国を 100 とし、使用なしの国を 0 にしたものである。相互の数値を比べにくい。項目によっては抜け落ちる国があるので、同じ国のデータを突き合わせるのが困難である。しかし 5 項目同時に数表をダウンロードできるので、全世界（全部の国・地域）でよく使われる 1 項目をいつも入れておくと、他の 4 語とそれとの相対的な数値が数表になり、安定した数値が得られる。ここでは anime を、常に 5 項目の一つとした。細かい技法については、井上（2012b）などを参照されたい。これまで 100 語以上の外行語について数表をダウンロードし、まとめて行列の形の数表を作成した³。いくつかは既発表である（井上 2012c, 2013b, Inoue 2012a, 2013d）。今回は行の合計値と列の合計値を算出して、よく使われる外行語の順番に並べ、また外行語をよく使う国の順番に並べた。

Google Trends のデータは、Google 検索に使われた語の集積に基づく。国家と時期（日）による集計結果がグラフ（および集計表）として出力されるが、計算方法などはブラックボックスである⁴。国家・地域ごとの集計だから、言語、文化、宗教とは食い違う面もある。

以下の考察では、因子が（非）言語的要因とどんな関係を示すかにも注目する。結論を先取りしてしまうと、経済要因が大きく働き、地理的近接効果も作用し、大陸（州）ごとにまとまった傾向を見せる。しかし同一言語を公用語とする国家が固まる傾向も見せる。宗教的要素の働きは観察されなかった。

Google Ngram Viewer を開発した Michel 他は *culturomics* という歴史的な研究分野を提唱したが（Michel et al. 2010）、Google Trends のデータは、地理的バージョンと言える。

3. 外行語と経済・文字の相関

外行語データに基づくこれまでの研究で、経済現象と関連がありそうな見通しはついていて、例えば井上（2012c）では国別の外行語使用率の順番のグラフに G8, ASEAN のメンバーかを記号で付して、外行語使用上位国と重なることを見出した。また公用語と文字が関わる見通しもあったので、英語を公用語にする国に印を付けた（井上 2012b）。ただ英語使用国の中でもアフリカ諸国に関しては外行語使用上位国にならないので、言語が順位に大きな影響を及ぼすとは見なさなかった。

今回改めて経済関係の連続的指標を求めて、外行語使用の連続的数値との相関をみることにした。指標としては各国の貿易輸出額を用いた⁵。ここでは外行語データのある国の

差も示される。数表ダウンロードには、右上の Settings の 3 点マークをクリックし、Download as CSV をクリックする。report.csv というファイルが送られてくる。この段階では Gmail に加入して、かつ使用している必要がある。

3 会社名も検索した。製品名が多くの人に普及していることから、外行語使用が少なくても検索からまれそうな国（例えばアフリカ諸国）を拾い上げるためもあった。さらに、普通名詞としての製品名と固有名詞としての会社名は区別しにくく、使用者の意識に常に相違があるわけではない。

4 グーグルで用いている技法については、例えば以下のサイトに示されている。

<https://www.google.co.jp/intl/en-US/policies/technologies/location-data/>

5 GDP を用いると、資源の多い小国家の数値がはねあがる。例えば一人あたりの数値だと、常識に反する数値になる。日本と他国との貿易に限った数値を探したが、主要国についての集計のみで、貿易額の小さな国家・地域を含めた集計は見当たらなかった。主要国については、井上（2012c）で示したように、経済要因と外行語使用に密接な関係が読み取れた。本稿では、世界全体を示す CIA 提供のデータを用いた。

みをグラフ化した。

図 1 に散布図を示す。横軸は外行語使用数で、145 国について 114 語の Google Trends の数値を合計したものである。縦軸は各国の貿易輸出額である。文字どおり桁違いで、下位の国家は数値が重なって国名が読み取れないので、対数を用いて表示した⁶。図 1 の国名によって考察する⁷。近似線の左上になるのは、中国・韓国を筆頭にする東アジア諸国のほか、西欧諸国とアメリカである。貿易額で予測するよりも外行語が少ない。それに対し右下には東南アジア諸国のほかに南米諸国が並ぶ。貿易額が少ない割に外行語が多い。以下の第 1 因子として目立った国でもあり、日系人が戦前から進出していた国でもある。西アジア諸国は近似線の間あたりにあまり離れずに分布する。アフリカ諸国は近似線の下端あたりにあまり離れずに分布する。最下位のマダガスカルは別格で、貿易額が少ないのに外行語が多い。第 6 節で論じるように木越（2004）では日本語観国際センサスの約 30 カ国の数値を近隣諸国に適用して全世界の数値を推計したが、図 1 によると、大きなずれはなさそうである。

(1) 近似直線のとおり、114 語の外行語使用度全体として、貿易額と比例関係を示した。縦軸に対数を使わないグラフ（本論文では略）によれば、カナダを通してアメリカに向かう近似直線に沿って各国が並ぶ。経済力の大きい国は同時に日本との交流も多い国なので、外行語が多くなると説明できる。経済は言語に影響を与える。中間項として文化交流なども設定できる。

(2) もう一つの要因がある。図 1 で中国・韓国などが近似線の左上になり、経済力が大きいのにその割に外行語使用は少ないという傾向を示した。一方近似線の右下（図の右中ほど）のシンガポール、フィリピン、インドネシア、マレーシアのような東南アジアのアルファベット国は、経済力は中程度なのに外行語使用は多い。

国によって 2012 年、2013 年のデータが混じる。今回は輸出額を用いたが輸入額のほうが外行語とよく対応する可能性がある。ただしグラフで全体を見渡すときには影響が小さい。輸出と輸入の額が大きく異なるような国家は少ないからである。また、もっと以前の数値が得られれば、グラフの近似がよくなる可能性がある。図 1 で左上に位置する中国や韓国は近年経済力（貿易額）が急に増えたが、昔は左の真ん中あたりに位置していたはずである。

6 そのために、近似直線を入れたが曲線として表示される。近似直線は $y=200000000x + 8000000000$ である。エクセル上では係数は各々 $2E+08$ 、 $8E+09$ と指数表記されて表示されるが、指数表記の + が加算の + と紛らわしいため削除してある。なお、指数表記の数値はエクセルのセルに入力し、表示形式を「標準」にすると通常の表記に変換できる。（査読者の補足による。）

7 文字が小さくて読み取りにくい、大きくすると重なって読み取れない、この大きさが妥協点である。2 文字か 3 文字の国名略号の使用も考えたが、用意した別表と照合する手間がかかるので、本論文では採用を控えた。

さらに詳しく見ると、近似線の左上には、アルファベットを使わず、漢字、ロシア文字、アラビア文字、インド文字、タイ文字などを使うアジア諸国がプロットされる⁸。貿易額がもっと下位の国・地域をみても、アルファベットを使わない国は左に寄る⁹。ここから Google 検索の外行語使用に文字が影響を与えると結論できる。非アルファベット国では Google 検索につづりとしてのアルファベット外行語を使わないので、数値が低くなったという説明である。以前から国家の順番の表で気づいていたことだが、今回の散布図の、実証的データにより確認できた。

(3) もう一つ、言語政策として外来語を規制し、または自国語要素を活かして造語するために外行語が少なくなったという説明もありうる。中国語、韓国語にあてはまる。英語排除の方針としたフランスにもあてはまり、線の左上に位置する。しかしハンガリーやバルト諸国のように自国語要素を活用する言語政策をとる国が線よりも右下になるので、日本語からの外行語については排除・置き換えが徹底しないようである。

4. 因子分析の手法

世界分布の大きな傾向はこれまでの地図の視察によりつかめていた。さらに細かく見ると多くの要因が働くことが予想されたので、因子分析を適用した。元データは 145 国家・地域 × 114 語の行列である。数の少ないセルは因子分析の全体像を乱すことがあるので、行列の合計値によってソートして、少数のセルを含む国家・地域と語を削除した。79 国家・地域 × 111 語の数値行列のデータに因子分析を適用した¹⁰。

プログラムパッケージ SPSS によって因子分析を施した。因子分析の中でも、様々の計算法があるが、データの性格および解析の目標から、繰り返し計算つきの主因子法により、バリマックス回転を施した¹¹。得られた Google Trends のデータの国・地域については、因子分析の「Q 技法 (QFACTOR)」(三好 1962) のように、地点 (ケース) と語形 (変数) の位置を入れかえて適用した。語の図は、井上・柳村 (2015) で論じた¹²。

初期の固有値が 1.0 以上に達した因子が 16 個得られたが、寄与率を考慮して、そのうち第 10 因子までを説明因子として採用した。第 11 因子以下は、後述するように個別的な

8 Japan も以下の図表に表れる。アルファベット以外を (国字として) 用いる国でもインターネット検索でアルファベットを用いる場合があり、その場合は Google Trends のデータとして扱われる。以下の考察でアジア諸国はアルファベット圏と異なった傾向を示す現象と関連して、用心すべきである。日本国内でも言語景観として monolingual (アルファベット単独) 表記が増えており (井上 2011b), tsunami, sushi のような表記は外行語の逆輸入の場合もある。Walkman をはじめ、日本語なのかの判別は今や困難になりつつある。

9 Google Trends の外行語使用がゼロの国・地域を入れた散布図も作ったが、左下に固まるだけで、新情報は得られなかった。つまり貿易額がある程度あるような国・地域では、外行語を使う。

10 計算から除外した 66 の国・地域は、例えば、以下のように大部分がアフリカや西アジアの小国家である。Afghanistan, Burkina Faso, Burundi, Benin, Congo, Eritrea, Ethiopia, Gambia, Guinea, Lesotho, Mali, Mauritania, Malawi, Niger, Papua New Guinea, Rwanda, Swaziland, Togo, Tajikistan, Turkmenistan.

計算から除外した語は、以下のとおり。hinoshi, mebos, shinkansei (ミスタイプ)

11 16 回の反復で回転が収束した。各因子の寄与率は表 1 のとおりで、大きな説明力を有しないので、以下のグラフの第 9 因子以降の読み取りには一定の留保が必要である。

12 単語ごとの因子分析結果はすでに井上・柳村 (2015) として別に公表した。少数の語例によって全語または全国家のパターンが偏ることを避けるために (しかし一定の信頼度を保つために) 単語をまた国家を分析から削除したため、別論文にした。

関係を示すもので、再現されない因子になる恐れがあるため解説を加えない¹³。データの KMO 値は 0.716, バートレットの球面性検定は 0.1% 水準で有意であり、このデータが因子分析に適合することが確認された。各因子の固有値と寄与率を表 1 に示す。

表 1 因子の固有値と寄与率

	初期の 固有値	回転後の負荷量平方和		
		寄与率(%)	累積(%)	
因子	1	23.693	14.973	14.973
	2	6.328	10.063	25.036
	3	4.990	5.581	30.617
	4	3.899	5.465	36.083
	5	3.148	4.929	41.011
	6	2.619	4.797	45.809
	7	2.223	4.527	50.335
	8	1.993	4.090	54.425
	9	1.687	2.413	56.838
	10	1.637	2.241	59.078
	11	1.445	1.953	61.031
	12	1.295	1.919	62.950
	13	1.190	1.820	64.770
	14	1.145	1.699	66.469
	15	1.118	1.433	67.902
	16	1.084	1.120	69.022

5. 国ごとの因子分析の結果

以下では国ごとの分析を、2 因子ずつの散布図として考察する。なお以下の図の読み取りを補足する資料として、各図に表示する数値を因子行列の形で本稿末尾の資料 1, 2 に示す。国・地域ごとの各因子の詳細な値についてはこれを参照してほしい。まず図 2 で第 1, 2 因子組み合わせて示す。グラフ表示に際しては、記号の位置を中央にして、数値マーカーに重ねた¹⁴。図 2 の視察により読み取りを行う。国・地域の性格に一定の傾向が見られた。国家分類の記号は以下のとおりである。

- 主要英語国, ●西ヨーロッパ, ○東ヨーロッパ,
▲東アジア, △西アジア, ×アフリカ, +中南米。

5.1 第 1 因子 中南米因子 (現地定着・食品因子)

図の右、横軸第 1 因子の値が特に大きい語として、右側に + 中南米諸国が分離された。唯一の例外は、+ ブラジルが真ん中に位置することである¹⁵。▲東南アジア (フィリピン

13 この種の研究は他に見つからなかったもので、どの程度の因子が世界の分布に関係するか予測できなかった。今回グラフを考察して見当がついた。

14 ■●○▲△×+ の記号は、国名ラベルの文字列の先頭に入れた。この文字列をダウンロード済みの "XY Chart Labels" によって、通常のエクセル散布図に貼り付けた。そのため、長い国名の場合、記号の位置が左にずれることに注意。

15 (0.2, 0.2) 付近。この図はもっと細かい字にしていたが、読み取れないので大きくした。

など), △西アジア, ×アフリカ (マダガスカル・エジプト) などが中間になる。これに対して第 1 因子の値が小さい国は, 図上端の●西欧, ○東欧や, 原点付近の▲東南アジア・オセアニア諸国である。第 1 因子は「中南米因子」と言える。

なお井上・柳村 (2015) 所載の 125 国家・地域×59 語の数値行列によれば, 語の因子分析による第 1 因子は, 「現地定着・食品因子」と名付けられ, 食品 (tofu, sushi, teriyaki, wasabi, sake, mochi, tempura), 伝統文化に関わる語 (futon, haiku, kimono), 現地化した語 (hobo = 浮浪者) を含む。日系人による密接借用 *intimate borrowing* の可能性がある (Bloomfield 1933)。図 1 では, 貿易額が少ない割に外行語が多い国で, 戦前からの日系人居住のあった地域に重なる。なお井上 (2013b) によれば, 食品は移住者によっていち早く近隣住民に普及する語彙とみられる。

これまで南米因子の働きが目立たなかったのは, 世界地図上で, 欧米と東南アジアにおける大きな使用率が際立っていたからである。個々のグーグルマップに戻ると, ブラジル・アルゼンチンなどがかなり大きな数値 (濃い色) を示している (Ogawa 2001)。井上・柳村 (2015) に掲げた世界地図にも示されている。地図の個別的・主観的視察の限界を示し, かつ今回の計量的手法の適用成果を示すとよい。

5.2 第 2 因子 西欧因子 (現代若者文化因子)

図の上, 縦軸第 2 因子の値の大きい国としては, ●西欧諸国が固まる。また○東ヨーロッパ (中欧) のチェコ, ポーランド, ルーマニア, ハンガリー, △西アジアのレバノンも含まれる。第 2 因子の値の小さい国の大部分は, 第 1 因子の大きいグループでもある。第 2 因子は「西欧因子」と考えられる。

図 2 の上端, 第 2 因子の値の大きいドイツ, オーストリア, (4 公用語の) スイスは, ドイツ語圏とも言える。ここに近くプロットされたのは, ドイツ語の影響の大きい国で, 北欧のデンマーク, ノルウェー, フィンランドなど, 中欧に属するチェコ, ポーランド, ハンガリー, スロバキアである。やや離れてスロバキアに重なってフランスとイタリアもプロットされる¹⁶。ロマンス語圏とまとめられ, (ロマンス語圏の) ルーマニアと (フランス委任統治領を経験した) レバノンの第 2 因子の大きさもこれによって説明される可能性がある。

図 2 の分布は, 左下から見ると, 日本を中央においたときの世界地図の東西配置に似る (アフリカを下に置ければ, さらに似る)。ただし地理的要因だけが働くわけではない。前述のように中南米の中でもブラジルが離れるのは¹⁷, ポルトガル語が公用語のためと考えられる。一方ヨーロッパのポルトガルはブラジルに近くプロットされ, スペインもその原点寄りにプロットされる。中南米のスペイン語は, 方言分化が進んでいないとも言えるし, 外行語の受容に地域差が大きいとも言える (上田 1987, 2013)。なお (メキシコとアルゼンチンに近いことからいっても) 中米と南米を分ける傾向は読み取れない。

以上のように, 第 1, 2 因子は一見中南米と西欧という地理的近接効果による外行語受容の差を示すとみられたが, 細かく見ると言語的要因に支配されていることが読み取れた。第 1 因子は正確には「中南米スペイン語圏因子」と呼ぶべきである。

16 (0.1, 0.4) 付近。

17 (0.2, 0.2) の近く。

なお語の因子分析で第 2 因子の値の大きいものには現代日本文化, pop カルチャーを反映する新語が多い. Otaku, anime, manga, pokemon, walkman, karaoke, doraemon, kaizen などである. 語ごとの分析では, 第 2 因子を「現代若者文化因子」と名付けた. 文化借用 cultural borrowing だが, 英語をはじめとする西欧語からの借用語と性格が違う (Bloomfield 1933). 第 2 因子の値の大きい soy, soya (しょう油) は, オランダ語を経て英語などに入った例外的な語で (原口他 1998, 井上・柳村 2015), 肉類へのソースとして戦後世界に現物が普及した. 「現代若者文化因子」に近いのは興味深い.

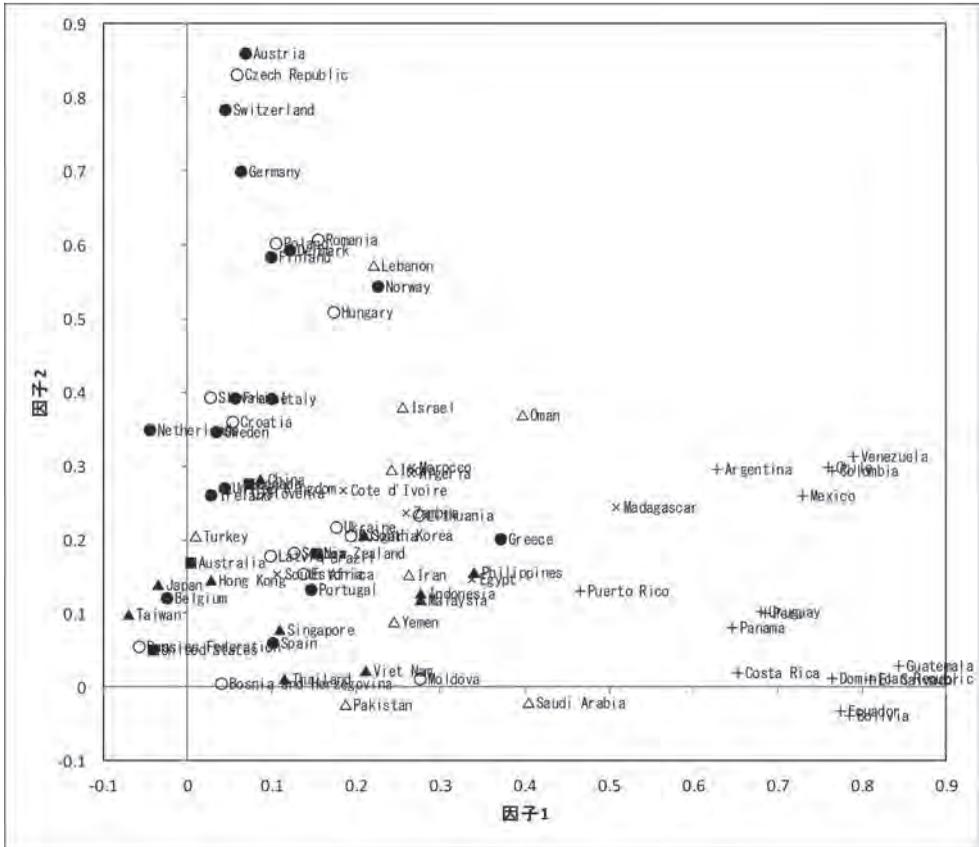


図 2 国ごとの分析 第 1,2 因子

5.3 第 3 因子 ヨーロッパ周辺因子

図 3 に第 3, 4 因子を組み合わせ示す. 図の右, 第 3 因子の値が大きい国は, △西アジアのレバノン, イエメン, イスラエル, オマーンである. また×アフリカのザンビア, アルジェリア, エジプトも大きい値で, 中位のサウディアラビア, インド, トルコに連なる. この中にバルト三国を筆頭にした○東欧とギリシャが混じる. 地理的に西ヨーロッパの南と西に連なる周辺国として共通性を持つ. アフリカの中でも南アフリカ, マダガスカル, モロッコ, コートディボアールは下の真ん中あたりにやや離れる. 「ヨーロッパ周辺

この2カ国でまったく違うからである。言語も文字も宗教も多様な国がまとまった傾向を見せるのは、文化的な共通性、または経済的交流の大きさによるのだろう。東アジアは、Google Trends の世界地図で外行語を用いる3大地域の一つである。また、GDPが高かつ（衛星写真「夜の地球 Earth Lights」によれば）夜も明るい地域である。経済と明るさと外行語使用には、共通の背景があると思われる（井上 2015c）。

5.5 第5因子 東アジア・ヨーロッパ周辺因子

図4に第5,6因子を組み合わせて示す。横軸第5因子の値が大きい国として、イラン、ウクライナ、コートディボワールと中国・台湾という地理的に離れた国が固まった。第5因子は「東アジア・ヨーロッパ周辺因子」と名付ける。他の多くも、アルファベット以外の（民族）文字を使う国である。中国、台湾、韓国という▲東アジアの旧漢字圏、イラン、イスラエル、ギリシャ、ウクライナを含み、中間にレバノン、イエメンが位置する。地理的に離れているコートディボワールが例外である。ただし、原点付近に固まる諸国家にも東南アジアや西アジアのように、民族文字を使うものが混じる。これらの国で、アルファベットつづりを使ってGoogle検索をするときには、その単語について偏りが出るのである。その国の言語・文字で完全に自国語と同化した語は、アルファベットによるGoogle検索はなされない。まったく知られない単語は検索されない。従って、中間的な位置にある外来語が、これらの国に多くなるのだろう。

図1でみたように、中国、台湾、韓国は貿易額が大きい割に外行語が少なかった。一般的原则として自国語要素に置き換える国でもあり、かつての植民地政策に反発して純化政策をとった国でもある。網の目をくぐるようにいくつかの外行語が（Google検索で）使われたのだろう。

5.6 第6因子 英語圏因子

縦軸第6因子の値が大きい国として、英語を公用語にする■カナダ、アメリカ、英国、オーストラリアが固まった。ただし、ニュージーランドは（値が大きい）他の西欧諸国にまぎれる位置にある。英語を使う、昔のイギリスの影響の濃い国が固まったわけである。原点付近に離れるが、●西欧のアイランドと、オランダ、デンマーク、ノルウェイが、やや高い値を取る。ゲルマン系言語という共通性がある。言語的に違うがフィンランドも近い値である（ただしゲルマン系言語のスウェーデンは離れる）。英語を公用語（の一つ）にしている香港、シンガポール、マレーシア、南アフリカは、近くに位置する。

第6因子マイナス側には、▲東南アジア、+南米の一部、トルコ、ポルトガル、ロシアなどがある。いずれも英語直輸入に抵抗を示す国である。第6因子を「英語圏因子」と名付けよう。ただし、このグラフに表れない国にも注意を払うべきである。アフリカの英語を公用語（の一つ）にしている国の中には、語数が少ないために削除された国もあるし、そもそもここでの外行語の使用がゼロであるために登場しなかった国もある。元のデータ行列を使用率によってソートしたときにすでに表れていた傾向で（井上 2013a）、外行語使用が際立つ国は英語を公用語にしている傾向があるが、アフリカやインドにはあてはまらない。

第6因子になってはじめて言語的要因が表に出た。図4の第5,6因子では、地理的条件よりも文字と同系言語の条件が働くさまが窺えた。単純な計算でなく、因子分析にかけても、同様な関係が表れたわけである。

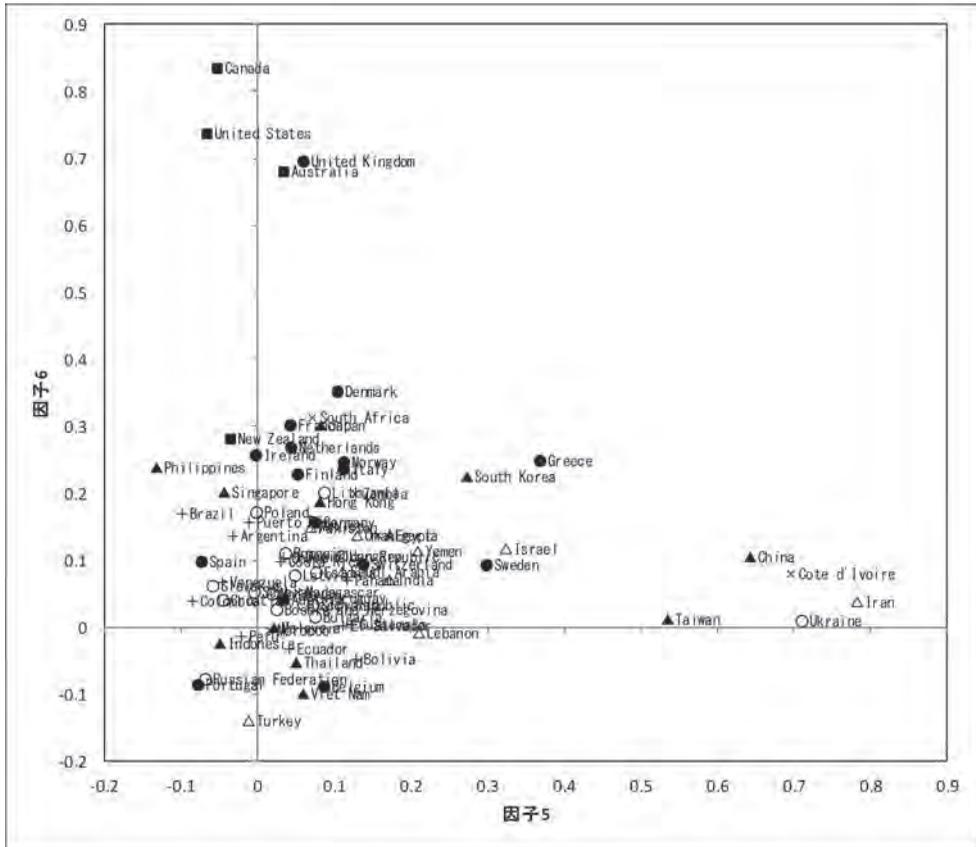


図4 国ごとの分析 第5,6因子

5.7 第7因子 西欧周辺因子

図5に第7,8因子を組み合わせて示す。横軸第7因子の値が大きい国は、ポルトガル、スペイン、アイルランド、スロバキアという●西欧周辺国と、ニュージーランドである。この因子では、大きいマイナス値をとる国が目立つ。オマーン、ザンビア、マダガスカル、イエメンという×アフリカと周辺国である。第7因子は「西欧周辺因子」と名付けうる。

5.8 第8因子 東欧因子

縦軸第8因子の値が大きい国は、○東欧4カ国の、セルビア、ボスニアヘルツェゴビナ、クロアチア、スロベニアで、かつてのユーゴスラビア、言語的にはかつてセルボクロアート語を公用語にしていた国である。外行語に共通性があるのは当然と言える。アルファベット（ローマ字）を使うかキリル（ロシア）文字を使うかの違いがあるが、造成言語として、今相互の違いを際立たせようとしている。この因子でマイナス値をとるのは、日本、タイをはじめとする▲東アジアの国などである。第8因子は「東欧因子」と名付けうる。

図5の第7,8因子では、地理的条件も言語的条件も、狭い地域に関するものになった。世界を大きく分割するような因子は地理的には第4因子までで、言語的な違いに関わるも

のの大規模な違いは第 5, 6 因子までと考えられる.

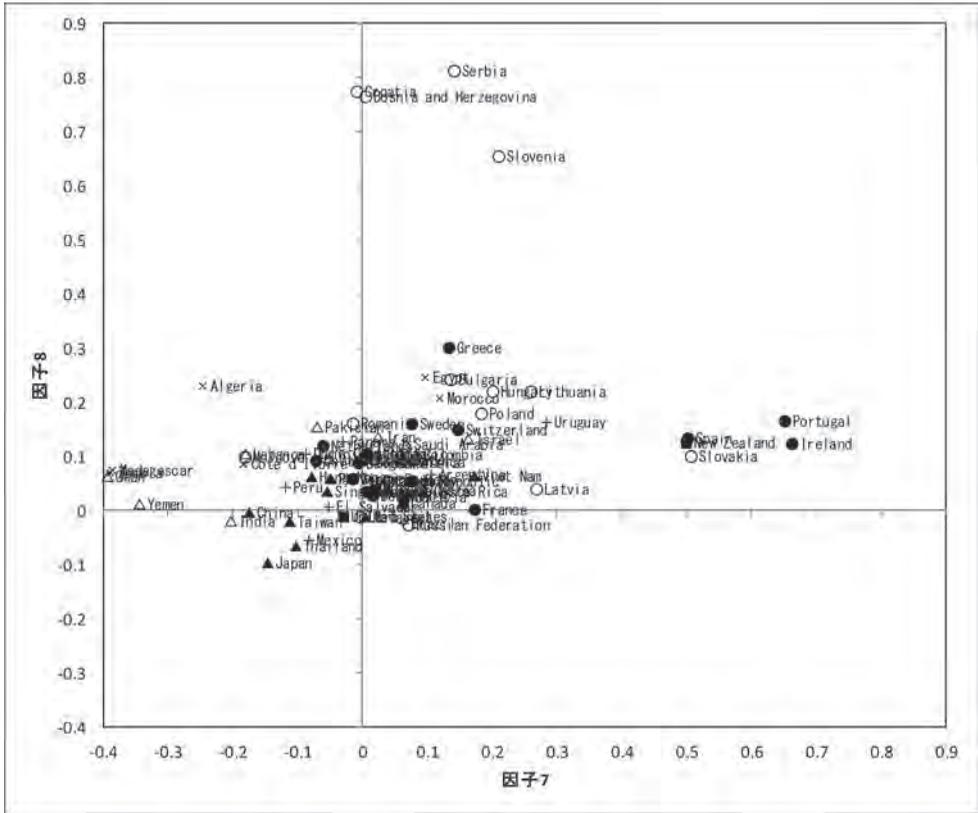


図5 国ごとの分析 第7,8因子

5.9 第9因子 親日国因子

図6に第9,10因子を組み合わせて示す。横軸第9因子の値が大きい国は、日本を筆頭にブラジル、●西欧のフランス、イタリアである。日本国内のアルファベット表記語と使用状況と似ているわけである。第9因子を「親日国因子」と名付けよう。中央の塊のうちのメキシコ、ポーランド、スペインを含めるとヨーロッパ周辺国として共通性を持つ。アルファベットによるグーグル検索が特徴を發揮し、他国と共通性を示したわけである。

5.10 第10因子 東欧・中米因子

縦軸第10因子の値が大きい国は、○ブルガリア、+中米のパナマ、プエルトリコである。離れてスロバキア、スロベニア、ルーマニア、リトアニアという○東欧が分布する。ギリシャが東欧に近いのは地理的に見て納得できる。西欧の周辺国の共通性を再び拾い出したのだろう。第10因子は「東欧・中米因子」といえる。このあたりになると、個別的・偶然的要素が多くなる。個々の外行語の使用率が強く働いたためと見られる¹⁹。

19 この結論によれば、第9因子以下を散布図で示す意義が薄れ、結果を数表で示してもいいことになる。

因子としては、第 11 因子以下も出てくるが、個別的な関係を示すもので、分析手法や単語の種類を変えただけで再現されない因子になる恐れがある。これ以降の因子には解説を加えない。数年後にデータを取り直して比較することも考えられる。

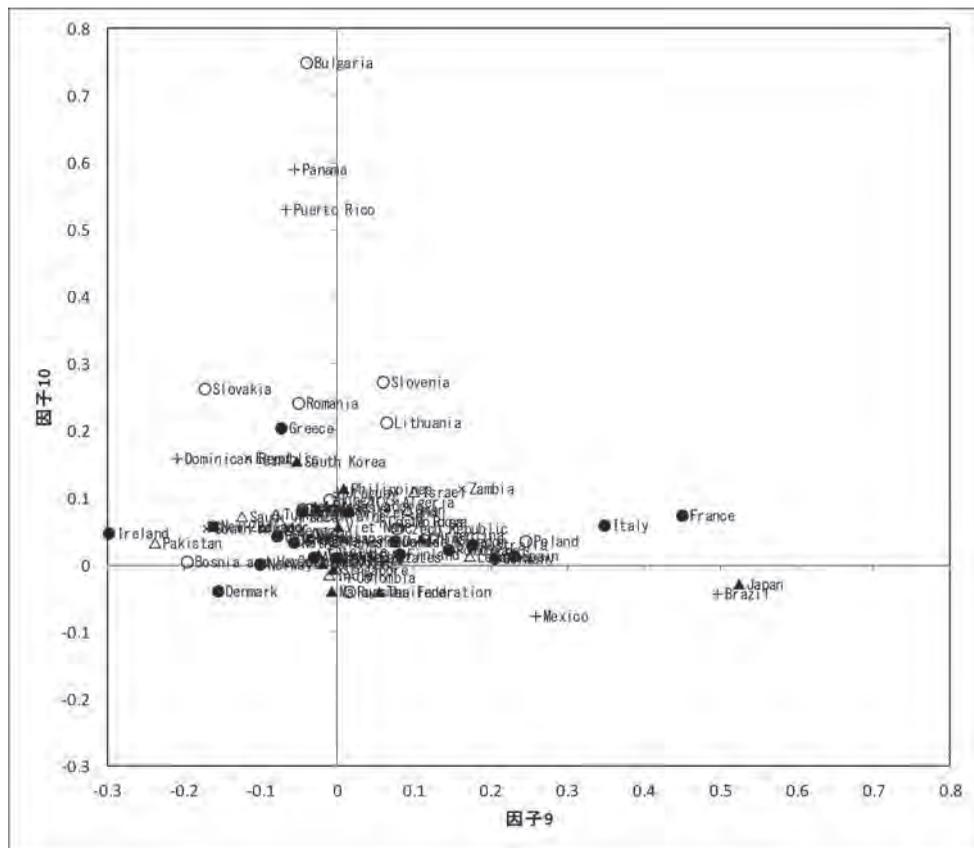


図 6 国ごとの分析 第 9,10 因子

6. 考察 世界諸言語の総合

6.1 外行語世界分布の位置付け

以上の因子分析の結果は、これまでの外行語研究の発展であり、総まとめの面も持つ。外行語に関しては、これまで国家・地域の行のソートと、単語の列のソートの結果をそれぞれ 1 次元のグラフとして考察した。順番によって大きな傾向を把握できた。さらに井上・柳村 (2015) 所載の 125 国家・地域 × 59 語の数値行列によれば、行のソート、列のソートを重ねるだけで、それぞれが 2 典型に分かれそうな傾向が読み取れた。今回の因子分析により、予想とかなり一致する結果が得られた (南米のまとまりは目立たなかった

とはいえこれが分かったのは散布図を使ったからである。数表に比べて、視覚化によって得られる新たな情報は多い。

が)。1次元的な見方を脱して、多次元的な見方にたどりついたわけである。

多くの世界地図を通じて、西欧、アメリカ、東南アジアが、外行語が多い地域として現れていた。今回因子分析を施したところ、これまで目立たなかった南米が第1因子として登場し、西欧は第2因子だった。東南アジアは第4因子、アメリカは第6因子として立ち現れた。地球規模の大きな傾向は存在するが、個々の国や語によって、事情が違ふことの反映である。語は独自の歴史を持つ (*Each word has its own history*) という、方言地理学・言語地理学・言語史のテーゼは、半分は成り立つというべきであろう。

ただし本稿図1に提示した経済要素は、外行語の多次元性を取り入れていない。言語外の要素としては、(海外における日本語の)言語景観について「人口」「経済」「文化度≡情報量」という3大要素を指摘した(井上 2011b, 2015c)。日本語言語景観と外行語使用には平行性が見られるので、将来はこれら全体をカバーする多変量解析に向かうべきである。いくつかの要素の世界全体の数値は(過去のものも含めて)把握可能なので、将来の課題としておく。

Bloomfield (1933) の古典的な分類によると、借用は文化借用 *cultural borrowing*、密接借用 *intimate borrowing*、方言借用 *dialect borrowing* に分けられる²⁰。文化借用と密接借用の区別は、1言語内の借用語(つまり外来語)をみた場合に、原語別にア priori に区別できる。借用に関わった言語の社会言語学的地位の上下(HかLか)または威信 *prestige* と関係する。これは個々の借用語の意味分野にも関係する。知的・抽象的分野には文化借用が見られ、物質的・具体的分野には密接借用が見られる。なお別の基準で語彙体系どうしの対照を行った場合には、下位 *Low* 言語からの借用が、受け入れ言語の欠落語(提供言語の特有語)でよく観察される(井上 2000a)。これは方言相互間、共通語と方言の間でも典型的に起こる。借用におけるこのプロセスは、文化人類学などでいう選択的受容 *selective adoption* に相当し、方言学で言うフィルター作用も同様の現象の言語面と言える。

これは、多言語について日本語からの外行語をみる場合にも、指摘できる。本稿の国家別のまとめ(南米、東アジア、欧米など)は、別稿で扱う単語別の因子分析と、コインの裏表の関係にある。言語の威信でいえば、日本語では英語をはじめとする欧米の諸言語に対してLの位置にあった。逆に東アジアにおいてはHの位置にあった。

例えば最近データが整備されつつある南太平洋(かつての南洋、ミクロネシアを含む)における外来語彙については(ロング 2015)、かなりが日本の信託統治、南洋庁時代の密接借用と考えられる。しかし日本語教育や教科書を通じて入った文化借用もありうる。アメリカ英語ハワイ方言(現地では *Pidgin* と呼ぶ)における外行語は、密接借用の典型である。語彙集を見ても、選択的受容が典型的に働き、日本語の特有語(英語の欠落語)に多い。ただしこの視点からの考察は、本稿のデータからは不十分にしかできない。なぜなら調査語彙の選定段階で英語への外行語の数が多く、現代・最近の世界各国への外行語は少ししか入れていないからである(とはいえ会社名や商品名も入れたために、日本と縁が薄く、借用が少ない国での借用の初期状況も把握できた)。中国語(小林 2015)、韓国語(ヤン 2007)をはじめ、浄化政策ともからんで借用語(音訳語)でなく意識語(漢字語)を頻用する言語については、補助的な資料を得る必要がある(荒井他 2013, 小林 2015)。

20 邦訳(三宅・日野 1962)では文化的借用、親密的借用、方言的借用と訳している。

またアジア諸国では国字（世界的にみれば地域限定の民族文字）が今やインターネットの世界でも盛んに使われるので、多くの文字における外行語を、総合して検索することが理想である。

外行語研究の直接の出発点は「日本語観国際センサス」の4語（キモノ、サクラ、ショージン、スキヤキ）の現地面接世論調査だった（国立国語研究所 2002）。日本語の地位などについては世界約 30 カ国の数値を世界地図にしてみても、地理的近接効果と経済的要素が働くという傾向が読み取れた。さらに約 30 カ国の数値を用い、人口を掛け算して、実際の比率を推計した。その後近隣諸国に同じ数値をあてはめて、世界全体の比率を推計した（木越 2004）。外行語 4 語の使用率についても、世界地図の分布をみると日本語の地位（人気度など）と比例する傾向が見えた（井上 2000b, 2001）。今回の図 1 の結果によると、大陸ごとにまとまる傾向がみられ、かつ近隣諸国は外行語使用率に類似がみられる。「日本語観国際センサス」による木越（2004）の推計方法に間違いがないことを確かめえた。そもそも世論調査が実施不可能なアフリカ諸国に、ナイジェリアなどの世論調査が可能な国の数値をあてはめることは心配だったが、今回の集計によると、アフリカ諸国のほとんどは外行語使用率が大変低い。また貿易額も小さいので、計算から除外した国も多い。中国などに比べると人口からいっても経済力からいっても、世界全体の数値に与える影響は少ない。

本稿のようなインターネット検索による研究は、従来の辞書の立項・収録状況などに基づく研究に比べて格段にすぐれている。辞書の研究では実際の使用頻度数は分からない。しかしインターネット検索によれば、使用頻度数が示される。「日本語観国際センサス」では世界約 30 カ国について、世論調査の形で4語の使用（意識）を尋ねた。この2種類の結果を突き合わせたところ、食い違いはあったが、合理的に説明できるもので、相互に信頼度の高さを補強するものだった。世界各国の辞書や借用語研究書を漁っても、「日本語観国際センサス」で得られた使用状況を反映するような結果を得ることは困難である。インターネット情報については、玉石混交であることや手法がブラックボックスであることから、研究材料として疑問を呈する向きもあるが、今回の一連の調査からは、得られる成果が大きかった。

なお「日本語観国際センサス」のデータでは、国ごとに使用率の性差・年齢差なども分かる。これを手がかりに例えば *Shogun* がテレビ番組や映画のおかげで広がったのか、*Sukiyaki* が歌（上を向いて歩こう）のおかげで広がったのかが分かる。加えて、どの国の線も壮年層がピークになり、高年齢層と 10 代は使用率が低いことが分かった。日本国内の文化庁他の世論調査においても、10 代は借用語の導入について抵抗感を持たないが、実際には知らないし、使いこなせない。敬語と同様に「成人後採用」が観察されるのである（井上 2015a）。類似の現象はパラオ語における外行語のうち「経済」や「大統領」にあたる語が若い世代に使われないことでも観察される。衰退の道を歩んでいる可能性もあるが、10 代の社会的場面の狭さ、言語能力の低さに基づくとも考えられる。年齢差はインターネット（検索）使用者の年齢差とも関連するので、データの読み取りには注意が必要である。

6.2 世界の借用語の研究法

本研究は、世界の諸言語を総合的に整理・分類する試みの中に位置づけられる。類型論

的研究が盛んになっているが、文法類型・音韻類型と違って、語彙は文化的要素が影響して個別言語的な要因が働き、扱いにくい。基礎語彙などに基づく系統論を離れると、全言語を平等に扱えない。しかし借用語なら可能である。Weinreich (1953) ですでに論じられているように、新概念の導入に際して音形を導入して借用語を用いる以外に、自国語の要素で意識することもある。中間的なものとして元の言語の語構成や発想を生かして借用翻訳 *loan translation* を行うこともある。音韻の制約によって語形が元の言語と離れることもある。文法的制約によって性・数・格の語形変化がともなうこともある。周辺の意味領域の語彙体系に左右される点もある。文字の類型（表音文字か、漢字のような表意文字か）も影響を与える。検索のときにアルファベットを使うかによっても、Google Trends の結果は左右される。外来文化受容に関してゆるやかか、純化主義かの違いもある。

この論文では、以上のような諸条件を超越する形で、日本語からの借用現象を「外行語」と名付けて、計量的に扱った。同様の研究は他の言語から広がった借用語についても可能である。一言語内に入ってきた「外来語」については、どの言語でも比較的簡単に研究ができ、膨大な蓄積がある。しかしその逆方向、ある言語から世界の諸言語にどう広がったかの研究は、これまで不可能に近かった。

Google Trends の元データは Google 検索に使われた語である。その内部はブラックボックスだが、恐らく日常の言語使用をかなり反映するものだろう。またインターネットで行きかっているビッグデータ、膨大な言語情報のエッセンスとっていい。文法的要素や基礎語彙のように、日常の話しことばと書きことばでよく使われるのに検索されにくい語とくらべると、性格の違いは否定できない。他方の極端としての術語や固有名詞とも、検索に使われる頻度が違う。しかし文化語彙として *culturomics* の研究の発展に寄与することは大きい (Michel et al.2010)。

大局的にとらえると、借用語の使用率は面積あたりの GDP の世界分布と比例する。また衛星写真「夜の地球 Earth Lights」で夜も明るい地域は、様々な言語からの借用語が多い傾向がある (井上 2015c)。ただしフランス語のように言語政策として英語借用語を排除して自国語要素で新たな造語を試みた言語については、隣接ロマンス語圏（スペイン語やポルトガル語）や旧植民地にまで影響を及ぼすことがあり、欧米の他にアフリカが際立つこともある。しかし多くの言語の外行語を総合的に集計し、相互に扱えば、大きな傾向が浮かび出ると予想される。世界（地図）を単語の借用関係に基づいていくつかの言語文化圏に分けることが可能になると予測できる。今回の結果をみる限り、宗教圏、民族文化圏にあたる地理的領域との関係は出て来なかった。

現在携帯電話、電子メールの普及により、いわゆるビッグデータとして音声や文字による発信・受信がどの地点で行われたかが分かるようになった。外国人の移動も把握できる。これらの情報と、外行語を含めた借用語全体の動向とを突き合わせる研究も将来可能になるだろう。言語接触がいつどのように行われるかが、大量のデータによって明らかになる。

さらに、Google Ngram Viewer を使うと (井上 2013c, 2015b, Inoue 2013d)、世界主要言語（英語、フランス語、ドイツ語、ヘブライ語、イタリア語、ロシア語、スペイン語、中国語）について 200 年間（または 500 年間）の (Google books における) 書籍中の語の出現率変遷を知ることができる（日本語は入っていない）。借用、翻訳借用、（民族主義的な浄化政策による）自国語への置き換えなどを、時系列として把握できる。いくつかの

外行語を試したところ、大航海時代または産業革命、近代国家形成の時期に初期的な借用語があり、20 世紀以降急激に借用が増えたという趨勢が読み取れた。本稿で扱った地理的（空間的）傾向に加えて、歴史的（時間的）傾向も、ビッグデータとして考察可能になった。言語現象の計量的研究の地平と時間軸が大幅に拡大したことになる。地理言語学と歴史言語学が接触し、相互に助け合う形になるだろう。その意味では、本稿は将来のグローバルな歴史地理学的言語研究、つまり国際借用語研究の序論とも位置付けうる。

7. 結論

本稿で扱ったのは日本語・英語をはじめとする外国語の影響が（実体計画としての）個々の借用語（つまりは外行語）に反映した現象である。その使用パターンが言語的要因と非言語的（地理的）要因の双方に支配されることが読み取れた。非言語的要因としては、文化交流に関わる要因以外に、経済的要因の働きの大きさが認められた。同様の現象は言語景観にも反映し（井上 2015b, 2015c）、（地位計画としての）外国語使用に関わり、また、同時に外来語の景観における露出と関わる。ただし民族文字の保持、翻訳における純化主義などの国家ごとの違いがあり（井上・包 2015）、言語外的な経済条件が直接に単純に作用するわけではない。さらに深い考察が必要だが、別稿にゆずる。

従来の言語的借用の研究は、主に基準になる 1 言語を出発点にするもので、研究のたやすさから、某言語に借用された外国語を扱うものだった。某言語が外国語にどう広がったかという研究はあったが、世界の諸言語への広がりを悉皆調査の形で実行することは不可能に近かった。インターネット情報を活用することでそれが可能になった。

これまで英語、スペイン語、ドイツ語、フランス語、イタリア語、ヒンディー語、韓国語、中国語、タイ語、ベトナム語からの貸付（Lending）のデータもあるが、別稿で一部を扱ったのみである（Inoue 2013d）。またハワイとアメリカ西海岸で日本語・韓国語・スペイン語などからの借用語を調査して報告した（井上 2013b）。個々の単語によって分布パターンが違うが、総合的に見ると、アメリカ、西ヨーロッパ、東南アジアに借用語が多く観察される。語の輸出国からの地理的近接効果と、語の輸入国の受容度によって説明できそうである。もっと語数を多くして、地図でなく数表としてダウンロードすれば、多くの単語の相互借用関係が分かるだろう。将来の課題である²¹。

これまでの論考では個別の語誌的記述でなく、多くの語彙の全体傾向をみようとして心がけた。ただし量的考察にとどまった。本論文で多変量解析法を適用することにより、世界の諸国家の言語を総合的に考察できた。とはいえ日本語を出発点にした輸出、貸付（Lending）の研究であり、他の言語から輸出された語は統合していない。

借用研究はともすると近代、現代のものに目を奪われる傾向があるが、古代以来普遍的な現象として借用があった。身近な日本語をはじめとして、東アジアでは中国語から圧倒的影響を受けつづけた。単語レベルにとどまらず、音韻にも文法にも及んだ（山田 1940）。他の諸言語でも同様で、言語史を記述するのに、歴史を通じての他言語からの影響を排除することはできない。近代の造語や翻訳借用を論じるにも、例えばヨーロッパにおけるギ

21 なお Google Trends のデータには、様々な制約、問題点があり、読み取りにおいて保留すべき点も多い。しかしここでは論を省くことにした。

リシャ語・ラテン語，東アジアにおける中国語と並んで，西アジアにおけるサンスクリットやペルシャ語の地位を認めるべきである。

またピジン・クレオール研究でも，個別言語間の接触を見るだけでなく，多くの言語を見渡して共通点・相違点を見出す研究が進んでいる。従来の研究対象は文字を通じて文化人の中で起こるような文化借用が主流だったが，話しことばを通じて庶民どうしで起こる密接借用にも視野が広がったわけである。

本稿では現代の日本語からの外行語について，方向性の視野を広げた。この新しい方向性は他の言語にも拡大可能である。従来の日本語と英語の関係を中心にしたアプローチに比べれば，これで地理的・空間的な広がりができる。さらに歴史的・時間的な広がりをつけることが可能になった。従来の言語別，言語系統別，地域別の羅列としての言語記述に加えて，最近是世界言語地図の形で各言語現象別に（個別に）地理的分布を見ることができ。これらを踏まえて世界の諸言語について，借用に関して「世界史」を記述する体制が整ったことになる。しかも歴史的背景を踏まえて，総合的に借用メカニズムを解明することも可能である。従来の借用語研究には個別の傾向をとらえる狭い研究が目立った。今後は大量の多様なデータを収集した上で，計量的手法を適用することが望まれる。

資料

図 2 から図 6 で示した国ごとの因子分析結果について、本稿で扱った第 1~10 因子の国ごとの数値を提示する。

資料 1 国ごとの分析の因子行列：■主要英語国，●西ヨーロッパ，○東ヨーロッパ

	因子									
	1	2	3	4	5	6	7	8	9	10
■ Australia	.044	.170	-.003	.062	.078	.681	.072	.041	.220	.031
■ Canada	.100	.277	.088	.167	-.023	.836	.097	.015	.105	.036
■ New Zealand	.201	.183	.121	.118	.017	.283	.561	.127	-.109	.059
■ United States	.016	.052	.059	.054	-.004	.738	.044	-.010	.062	.012
● Austria	.099	.861	.009	.003	.067	.042	.047	.106	.003	.012
● Belgium	.006	.122	.029	.279	.120	-.088	.032	.089	-.045	.044
● Denmark	.151	.594	.274	.131	.137	.353	.055	.029	-.122	-.038
● Finland	.129	.584	.100	.175	.086	.229	.045	.036	.115	.017
● France	.083	.393	-.012	.033	.072	.303	.208	.003	.480	.075
● Germany	.094	.701	-.045	-.063	.108	.158	.024	.060	.239	.011
● Greece	.398	.202	.412	.167	.398	.250	.168	.302	-.044	.204
● Ireland	.058	.261	-.005	-.022	.031	.258	.701	.125	-.265	.048
● Italy	.122	.392	-.042	.103	.136	.237	.105	.056	.373	.060
● Netherlands	.003	.350	.248	.054	.097	.269	.002	.121	-.003	.035
● Norway	.253	.544	.413	.261	.142	.248	.055	.036	-.071	.002
● Portugal	.181	.133	.075	-.005	-.039	-.085	.697	.167	.184	.023
● Spain	.123	.061	-.153	-.053	-.049	.099	.530	.135	.255	.014
● Sweden	.061	.347	.077	.059	.328	.094	.111	.161	-.016	.081
● Switzerland	.093	.784	.045	.148	.191	.095	.210	.150	.067	.080
● United Kingdom	.105	.271	.085	.113	.127	.697	.008	.093	.143	.036
○ Bosnia and Herzegovina	.136	.006	.156	-.071	.131	.027	.129	.766	-.090	.005
○ Bulgaria	.229	.206	.292	.020	.114	.016	.181	.243	-.001	.750
○ Croatia	.083	.361	.042	.144	-.012	.042	.030	.775	-.019	.038
○ Czech Republic	.120	.832	.256	.084	.128	.034	.119	.055	.145	.055
○ Estonia	.167	.155	.464	.046	.110	.082	.034	.107	-.012	.084
○ Hungary	.203	.510	.204	.163	.143	.107	.240	.221	.023	.098
○ Latvia	.125	.179	.423	-.017	.078	.078	.303	.041	.026	.021
○ Lithuania	.314	.234	.562	.046	.131	.202	.312	.221	.108	.212
○ Moldova	.306	.012	.455	.190	.052	.051	-.142	.100	.125	.066
○ Poland	.132	.603	.265	.046	.028	.173	.218	.180	.276	.036
○ Romania	.185	.608	.208	.141	.070	.111	.024	.162	-.017	.242
○ Russian Federation	.022	.056	.031	.135	.019	-.076	.174	-.025	.104	-.039
○ Serbia	.153	.183	.117	-.010	.025	.053	.176	.814	-.007	.084
○ Slovakia	.062	.394	.066	-.028	-.020	.063	.552	.102	-.134	.263
○ Slovenia	.119	.264	.235	.012	.111	.033	.255	.656	.099	.273
○ Ukraine	.206	.218	.226	.134	.743	.010	.036	-.011	.019	.075

資料 2 国ごとの分析の因子行列: ▲東アジア, △西アジア, ×アフリカ, +中南米

	因子									
	1	2	3	4	5	6	7	8	9	10
▲ China	.108	.282	.017	.343	.666	.106	-.147	-.003	.134	.040
▲ Hong Kong	.068	.146	.030	.809	.125	.188	-.027	.063	.025	.002
▲ Indonesia	.316	.128	.120	.374	-.005	-.023	.068	.033	.016	.087
▲ Japan	-.013	.140	-.030	.061	.107	.302	-.118	-.095	.548	-.028
▲ Malaysia	.311	.119	.103	.566	.060	.000	.050	-.011	.031	-.039
▲ Philippines	.388	.156	.077	.438	-.079	.239	.013	.060	.061	.114
▲ Singapore	.149	.079	.092	.805	.000	.203	-.003	.036	.039	-.006
▲ South Korea	.257	.207	.174	.526	.326	.226	.101	.038	.000	.155
▲ Taiwan	-.043	.099	-.052	.560	.564	.013	-.078	-.019	.005	.016
▲ Thailand	.150	.012	.041	.647	.089	-.052	-.058	-.065	.093	-.040
▲ Viet Nam	.246	.023	.014	.090	.098	-.098	.218	.065	.040	.058
△ India	.264	.296	.179	.190	.196	.072	-.175	-.018	.013	-.014
△ Iran	.279	.153	.304	.031	.800	.039	.046	.137	-.011	.048
△ Israel	.282	.380	.452	.253	.353	.118	.197	.133	.130	.110
△ Lebanon	.251	.573	.526	.068	.244	-.008	-.142	.105	.206	.013
△ Oman	.414	.370	.481	.217	.148	.138	-.372	.064	.111	.083
△ Pakistan	.222	-.023	.187	.130	.108	.150	-.025	.155	-.199	.034
△ Saudi Arabia	.458	-.020	.257	.235	.169	.084	.136	.123	-.066	.073
△ Turkey	.036	.205	.299	-.081	.018	-.139	-.016	.097	-.051	.076
△ Yemen	.267	.089	.465	.281	.233	.114	-.316	.013	-.037	.053
× Algeria	.295	.291	.468	.210	.190	.138	-.208	.232	.108	.095
× Cote d'Ivoire	.241	.268	.153	-.006	.758	.081	-.111	.088	.000	.006
× Egypt	.359	.147	.494	.019	.194	.138	.125	.247	-.092	.160
× Madagascar	.551	.245	.131	.082	.100	.054	-.330	.076	.010	.041
× Morocco	.296	.299	.176	.068	.049	-.004	.157	.209	.193	.023
× South Africa	.159	.154	.267	.073	.130	.314	.077	.091	-.112	.056
× Zambia	.286	.237	.512	.280	.157	.200	-.355	.071	.192	.115
+ Argentina	.668	.296	.047	.078	.012	.137	.158	.068	.164	.046
+ Bolivia	.815	-.038	.145	.060	.162	-.047	.107	.026	.042	.009
+ Brazil	.187	.177	.109	.100	-.069	.170	.097	-.022	.525	-.042
+ Chile	.781	.299	.059	.244	.058	.105	.042	.094	.053	.014
+ Colombia	.799	.294	.042	.170	-.046	.040	.128	.105	.054	-.018
+ Costa Rica	.696	.020	.202	.181	.078	.099	.154	.036	.104	.065
+ Dominican Republic	.876	.013	.129	-.042	.140	.106	.016	.106	-.121	.160
+ Ecuador	.804	-.032	.116	.062	.074	-.031	.091	.058	-.081	.057
+ El Salvador	.858	.011	.180	.032	.164	.004	.010	.008	.036	.089
+ Guatemala	.889	.030	.180	-.072	.166	.005	.072	.056	-.013	.040
+ Mexico	.756	.261	-.007	.195	.101	.152	-.048	-.053	.288	-.075
+ Panama	.673	.082	.126	-.028	.146	.071	.004	.129	-.026	.590
+ Peru	.702	.101	-.028	.202	-.003	-.012	-.097	.044	.063	.013
+ Puerto Rico	.514	.131	.167	.041	.042	.158	.053	.095	-.014	.531
+ Uruguay	.710	.103	.249	-.101	.123	.045	.323	.165	.034	.109
+ Venezuela	.829	.314	.059	.199	-.002	.068	.086	.090	.069	.036

付記

分担: Google Trends のデータは井上が用意し, 文章は井上が担当した. 多変量解析は, 柳村が鎌水兼貴の指導のもとに担当した.

井上関係の現段階文献情報は, 以下の Web サイトで公開された.

<http://innowayf.net/>

http://dictionary.sanseido-publ.co.jp/affil/person/inoue_fumio/index_eng.html

また以下の論文のいくつかも Web で読むことができる.

文献 (五十音順)

荒井智子・杜岱玲 (2013) 「台湾にある日本語」明海日本語 18 増刊, pp.345-353.

井上史雄 (2000a) 『東北方言の変遷』秋山書店.

井上史雄 (2000b) 『日本語の値段』大修館, pp.1-15.

井上史雄 (2001) 『日本語は生き残れるか—経済言語学の視点から』PHP 新書.

井上史雄 (2011a) 「Google 言語地理学入門」明海日本語 16, pp.43-52.

<http://www.urayasu.meikai.ac.jp/japanese/meikainihongo/16/043inoue.pdf>

井上史雄 (2011b) 『経済言語学論考—言語・方言・敬語の値打ち—』明治書院.

Inoue, Fumio (2012a) "Improvements in the sociolinguistic status of dialects as observed through linguistic landscapes" *Dialectologia* 8, pp.85-132.

http://www.publicacions.ub.edu/revistes/ejecuta_descarga.asp?codigo=725

井上史雄 (2012b) 「日本語世界進出のグーグル言語地理学—グーグルインサイトにみる外行語総合分布—」明海日本語 17, pp.29-42.

<http://www.urayasu.meikai.ac.jp/japanese/meikainihongo/17/inoue.pdf>

井上史雄 (2012c) 「日本語の世界進出—グーグルでみる外行語—」『外来語研究の新展開』おうふう, pp.97-111.

井上史雄 (2013a) 「日本語の国際進出」明海大学外国語学部論集 25, pp.1-15.

井上史雄 (2013b) 「ハワイと世界の外行語—質問紙とグーグルにみる日本語使用—」応用言語学研究 15, pp.93-103.

井上史雄 (2013c) 「外行語と外来語 (日本語の攻防)」日本語学 32 卷 12 月号, pp.70-79.

Inoue, Fumio (2013d) *Google Linguistic Geography and Worldwide Borrowings — Interlingual dialectology in Japanese and Spanish — (Special Lecture) Universitat de Barcelona.*

井上史雄 (2015a) 「敬語の成人後採用 — 岡崎敬語調査の「川の字」変化—」国語研プロジェクトレビュー5 卷 3 号, pp.98-107.

井上史雄 (2015b) 「経済言語学の考え方—東京オリンピックの言語景観—」日本語学 34 卷 5 月号, pp.2-13.

井上史雄 (2015c) 「言語景観の表層と深層—政治と経済—」日本学 (東国大学校) 40, pp.85-107.

井上史雄・柳村裕 (2015) 「外行語の世界分布地図 —Google Trends データの因子分析—」明海日本語 20, pp.1-10.

井上史雄・包聯群 (投稿中) 「内蒙古文字景観の社会言語学—文字の社会類型論—」社会

言語学.

上田博人 (1987) 『スペイン語の語彙の頻度と拡がり: 資料集 = Frecuencia y dispersión del vocabulario español』東京外国語大学語学研究所.

上田博人 (2013) 「広域スペイン語語彙バリエーション研究における新しい数量化の試み—日本語計量方言学の方法に学ぶ—」日本語・日本学研究 3, pp.59-90.

<http://repository.tufs.ac.jp/bitstream/10108/73005/1/jo000304.pdf>

エバンス・M・年恵 (1990) 『英語になった日本語—ことばに見るアメリカ最新事情』ジャパンタイムズ.

Ogawa, Patricia Kanae (2001) "Lenguas en Contacto: Estudio Linguístico de la Colectividad Japonesa en la Argentina". *Japones para hispanohablantes: IV Jornada de Lengua Japonesa para Hispanohablantes*. Buenos Aires, Argentina: Editorial Dunken, pp.53-62.

加藤秀俊・熊倉功夫 (1999) 『外国語になった日本語の事典』岩波書店.

木越勉 (2004) 「地理的距離からみた言語の使用価値」言語 33 (9), pp.80-83.

Cannon, Garland; Warren, Nicholas (1996) *The Japanese Contributions to the English Language, An Historical Dictionary*. Wiesbaden: Harrassowitz Verlag.

高名凱・劉正淡 (1958) 『現代漢語外来詞研究』文字改革出版社.

国立国語研究所 (2002) 『東アジアにおける日本語観国際センサス』国立国語研究所.

小林孝郎 (2015) 「外国語に借用された日本語」『ことばの借用』朝倉書店.

新プロ「日本語」総括班編 (1999) 『日本語観国際センサス 単純集計表 (暫定速報版)』国立国語研究所.

http://www.ninjal.ac.jp/archives/n_census/result/

孫常叙 (1957) 『漢語詞彙』吉林人民出版社.

沈国威 (2008) 『近代日中語彙交流史』笠間書院.

永島恭子 (2013) 「グーグルインサイトに見る外行語の分布」明海日本語 18 増刊, pp.337-343.

早川勇 (2003) 「英語に入った日本語語彙の初出年調査」日本語科学 13, pp.79-108.

早川勇 (2006) 『英語になった日本語』春風社.

原口庄輔・原口友子編訳 (1998) 『新「国際日本語」講座』洋販出版.

Bloomfield, Leonard (1933) *Language*. New York: Henry Holt and Company (三宅鴻・日野資純訳 (1962) 『言語』大修館).

彭広陸 (2013) 「中国語の新語に見られる日本語からの借用語」日本語学 32 卷 11 月号, pp.14-25.

Michel, Jean-Baptiste et al. (2010) Quantitative analysis of culture using millions of digitized books. *Science*. Vol.331, no.6014, pp.176-182. (Published online ahead of print: 12/16/2010).

三好稔編著 (1962) 『心理学と因子分析』誠信書房.

三輪卓爾 (1977) 「外行語の昨日と今日—海を渡った日本語—」言語生活 312, pp.49-57.

山田孝雄 (1940) 『国語の中における漢語の研究』宝文館.

梁敏鎬 (2007) 「外来語をめぐる意識に関する日韓対照研究」国語学研究 46 号, pp.73-85.

ロング, ダニエル (2015) 「『太平洋日本語借用語逆引き辞典』におけるハワイ英語を含む

意義」第 5 回接触言語学国際会議（首都大学東京）配布資料。
ロング, ダニエル・長門正大（2015）「ハワイ英語で使われている日本語起源借用語」『日
本語研究』35, pp.139-152.
Weinreich, Uriel（1953）*Languages in Contact*, New York: Mouton.

（2015 年 2 月 15 日受付・2015 年 6 月 7 日再受付）

Paper

Country-wise Factor Analysis of
Worldwide Distribution of “Lent Words”:
The Tendencies by Google Trends

INOUE Fumio (National Institute for Japanese Language and Linguistics)
YANAGIMURA Yu (National Institute for Japanese Language and Linguistics)

Abstract:

It is possible to draw world maps for “lent words” or words exported from Japanese into foreign countries using an internet search engine. A numerical data of Google Trends can be downloaded; factor analysis was applied after putting the data in order, and the general tendencies of the maps were grasped. United States, Western Europe and South East Asia stood out as the area where “lent words” are used much, but Latin America showed the tendency different from other countries as a state group. Data of Google Trends could elucidate global distribution of “lent words” which has been impossible in the past.

Keywords: lent word, loan word, factor analysis, Google Trends, geolinguistics