

新聞に多用される外来語の分析 一面種ごとの特化係数を用いて一

桐生りか, 山口昌也 (国立国語研究所)

1. はじめに

本研究は、新聞で使用される語彙の性質を明らかにするための一考察として、新聞で多用される外来語を面種との関係から分析するものである。具体的には、面種ごとに、その面に特化して現れる「特徴語」となる外来語を特化係数に基づいて求め、まずは、各面に占める特徴語の割合を比較する。その後、各面の特徴語を、特徴語となる面種の範囲が広い語か狭い語か、どのような意味分野の語であるか、のふたつの観点から整理し、各面種の性質と、特徴語となる外来語の性質の考察を試みる。

2. 調査対象

調査対象としては「CD-毎日新聞」1994～2003年(日外アソシエーツ)を使用し、まずは、以下の基準(正規表現)によって、カタカナ・アルファベット・数字・一部の記号を含む文字列を抽出した。

$$[0-9\text{ア}\cdot\text{ヴ}\text{a}\cdot\text{z}\text{A}\cdot\text{Z}][0-9\text{ア}\cdot\text{ヴ}\text{ー}\text{a}\cdot\text{z}\text{A}\cdot\text{Z}\text{.}\text{/}\cdot\text{:}]$$

$$* [0-9\text{ア}\cdot\text{ヴ}\text{ー}\text{a}\cdot\text{z}\text{A}\cdot\text{Z}]$$
 ※ただし、数字と \cdot だけからなる文字列は除く。

その後、抽出結果から、10年間の毎日新聞全体における総出現頻度が1000を超す文字列を取り出し、さらにそこから和語や漢語のカタカナ表記・アルファベット文字列・アルファベットまたは数字を含む文字列を除外した¹。その結果残った1144語を“新聞に多用される外来語(以下、「高頻度外来語」)”として以降の分析対象とする。

「CD-毎日新聞」東京版には、以下の16の面種が設けられている。また、各面種に含まれる高頻度外来語の語数は以下の図に示すとおりである。

1面	2面	3面	スポーツ	家庭
科学	解説	経済	芸能	国際
社説	総合	特集	読書	文化

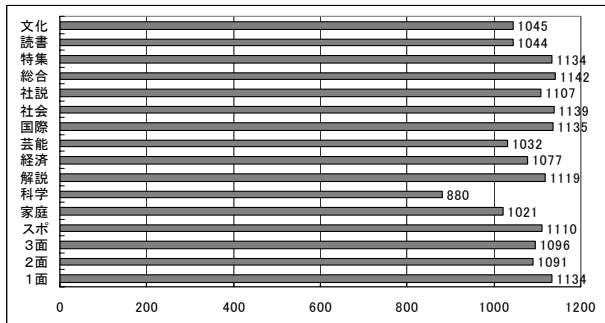


図1：高頻度外来語の面種別の異なり語数

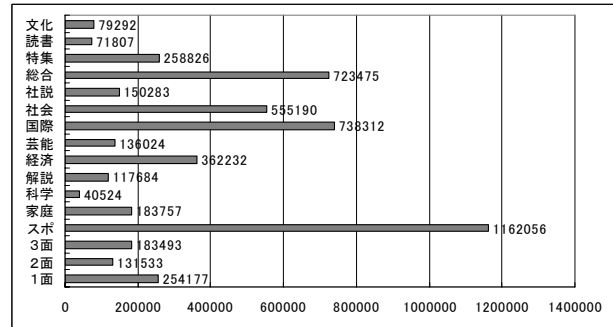


図2：高頻度外来語の面種別の延べ語数

3. 各面の特徴語の割合

3.1. 特化係数と特徴語

分析に先立ち、高頻度外来語の16面それぞれにおける特化係数を算出し、面種ごとに特徴語を求めた。

特化係数とは、ある語の新聞全体における構成比を標準とした場合に、ある面種における構成比が標準の何倍に相当するかを算出したもので、カタカナ語*i*の面*j*における構成比を p_{ij} とすると、特化係数 f_{ij} は、以下のように定義される。

$$f_{ij} = \frac{p_{ij}}{Q_i} \quad \text{※} Q_i \text{は16面合計における} i \text{の構成比。}$$

特化係数は1に近ければ近いほど、その面種におけるその語の構成比は標準的であり、1より小さければ小さいほど標準より小さく、1より大きければ大きいほど標準より大きい。本研究では、特化係数が標準値1を上回る語をその面の「特徴語」とし、考察を行うこととする。

3.2. 高頻度外来語に占める特徴語の割合

高頻度外来語を一般語と固有名詞に区別した上で、面種ごとに高頻度外来語全体に対する特徴語の割合を求めると、異なり語数では図3、延べ語数では図4のようになる。ここでは人名・地名・組織名と、記事のタイトルなどのうち一般用法を持たないと思われるもの(連載記事「おすすめヒット・ナウ」の「ヒト・ナウ」など)のみを固有名詞とし、その他の語はすべて一般語とした。なお、高頻度外来語全体1144語の内訳は、一般語770語(67.3%)、固有名詞374語(32.7%)である。

まずは、一般語・固有名詞を合わせた特徴語全体を見てみる。異なり語数では、特徴語が最も多いのは総合面で、高頻度外来語の45%近くを特徴語が占めている。次いで特集、3面、解説、社会、社説、1面の順に多い。反対に特徴語が最も少ないのは芸能面で、特徴語の割合は全体の23%に満たなかった。次いで国際、科学、経済、2面、スポーツの順となる。

¹ 固有名詞は語種を考慮しないものとし、和名であっても対象に含めた(ダイエーなど)。便宜上、外来語と呼ぶ。

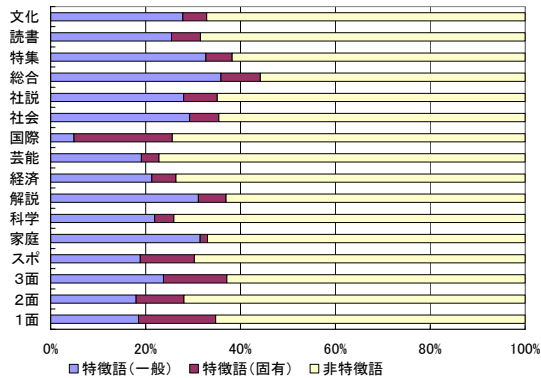


図3：面種ごとの特徴語の割合（異なり語数）

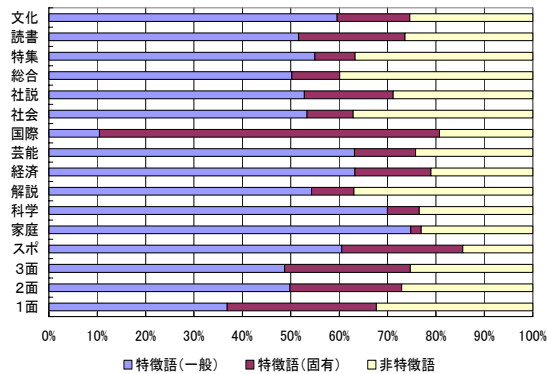


図4：面種ごとの特徴語の割合（延べ語数）

延べ語数を見ると、特徴語が占める割合が最も多いのはスポーツ面である。スポーツ面の特徴語は異なり語数では30%ほどと他面と比べてそれほど多くはないが、延べ語数では85%以上を特徴語が占めるという結果となった。次いで多いのは国際面で、異なり語数では特徴語は全体のわずか25%であるが、延べ語数では80%を占めている。科学や経済、芸能なども、異なり語数では少ないが延べ語数では多いという、同じ傾向を示した。反対に、特徴語が延べ語数に占める割合の低い方を見ると、最も低いのは総合面である。総合面は異なり語数では特徴語が全体の45%近くを占め最多であったが、延べ語数では60%でしかない。次いで低いのが社会、解説、特集であるが、これらの面も総合面同様、異なり語数における特徴語の割合は、他面と比べて高い面である。

相対的に、新聞の面種においては、異なり語数で特徴語の占める割合の多い面は延べ語数ではその割合があまり多くなく、異なり語数であまり多くない面は延べ語数では多くを占める、という傾向にある。例外としては、異なり語数・延べ語数ともに比較的少ない方に位置する2面と、反対に、両者とも比較的多い方に位置する3面が挙げられる。

次に、特徴語が一般語であるか固有名詞であるかという観点を加えて各面を見てみると、国際面における固有名詞の多さが際立っている。国際面は、特徴語の割合が異なり語数で全体の25%ほどと非常に少ないが、そのうち20%を固有名詞が占めている。この割合は、他面と比較しても格段に大きい。そして延

べ語数では、特徴語が全体の80%強を占めるうち、固有名詞である特徴語は実に70%に相当する。先に述べたとおり、高頻度外来語全体における一般語と固有名詞の割合は6.7対3.3であるが、特徴語において固有名詞の割合がそれよりも多い面は、国際のほかにも1面、スポーツ、2面、3面である。しかし国際以外の面は、高頻度外来語全体に占める固有名詞の特徴語の割合はそれぞれ異なり語数で10~20%、延べ語数で23~30%ほどであり、国際面がいかにか突出しているかが分かる。反対に特徴語に固有名詞が少ないのは家庭面であり、特徴語である固有名詞は、異なり語数で全体の1.5%、延べ語数では2%にも満たないという結果であった。

4. 特徴語の範囲

ひとつの語はひとつの面のみの特徴語であるとは限らず、複数の面の特徴語となり得る。そこで、高頻度外来語それぞれについて、それがいくつの面の特徴語であるのかを確認した。特徴語となった面の数をその語の「範囲」とすると、最大は「コピー」「メッセージ」の2語で範囲13であった。範囲1~13のそれぞれに該当する語数を、一般語・固有名詞に区別した上で図5に示す。範囲の平均値を求めると、一般語が5.8、固有名詞が3.8で、固有名詞の方が一般語に比べて小さい。そのことは図5からも見て取れる。

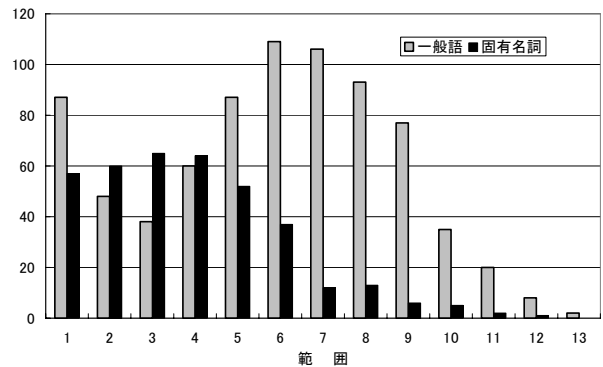


図5：特徴語の範囲

次に特徴語の範囲を面種ごとに割合で示したのが図6である。

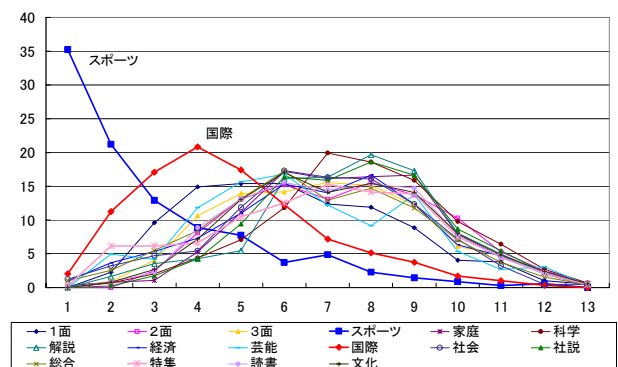


図6：面種ごとの特徴語の範囲

これを見ると、多くの面の特徴語が、山型に近い形の分布を示すのに対し、他面とまったく異なる分布を示したのがスポーツ面である。スポーツ面の特徴語には、範囲1、すなわちスポーツ面以外では特徴語とならない語が123語と圧倒的に多く、範囲の平均値は3.0であった。この123語という語数は、範囲1の語全体の85%以上に相当する。123語の中には外国人スポーツ選手の名前やチーム名などの固有名詞も45語含まれるが、残りの78語は一般語であり、スポーツ面の特徴語に、他面に現れない専門的な語がいくかに多いかがわかる。

他面についても面ごとに範囲の平均値を求めてみると、ほぼ全ての面が6～7.5の値を示すのに対し、ひとつだけ4.8と小さい値を示したのが国際面である。3章で述べたとおり、国際面の特徴語には固有名詞が圧倒的に多く、固有名詞の方が一般語に比べて範囲の平均値が小さいことを反映した結果と考えられる。

範囲1の語のほとんどがスポーツ面の特徴語であることは先に述べたが、スポーツ面以外の範囲1の語は、連載記事のタイトルやコラム名などとしてある面で継続的に使われた「ポシュット」「クロスロード」などの語や、「ナンバーズ²」など、ある面で定期的に掲載される記事の語、その他、国際面でのみ特徴語となった地名、経済面でのみ特徴語となった企業名であった。

反対に特徴語となる範囲の非常に広い語は、範囲13が前述の「コピー」「メッセージ」の2語、範囲12が「ギャップ」「ターゲット」「パターン」「バランス」「サラリーマン」「シベリア」「キーワード」「プログラム」「マニュアル」の9語である。「コピー」はスポーツ・芸能・国際以外の13面の、「メッセージ」はスポーツ、経済、国際以外の13面の特徴語となっている。範囲12の9語が特徴語となる面として比較的共通するのは、2面・3面・家庭・科学・解説・社説・総合・特集・読書・文化である。

5. 特徴語の意味分野

最後に、各面の特徴語の意味分野からの考察を試みる。考察に際しては、『分類語彙表』（国語研究所）を基準とし、各面の特徴語を分類番号の小数点第一位の数で整理した。小数点第一位の数は、整数で表される「体の類(1)・用の類(2)・相の類(3)」の間で統一された意味分野を表すので、たとえば、体・用・相それぞれにおける意味分野「抽象的關係」を表す1.1と2.1と3.1は区別せず、同じ「①抽象關係」としてまとめることとした。したがって、分類項目は、次の5つとなる。

- ①：抽象的關係 (1.1, 2.1, 3.1)
- ②：人間活動の主体 (1.2)
- ③：精神および行為 (1.3, 2.3, 3.3)
- ④：生産物および用具 (1.4)
- ⑤：自然物および自然現象 (1.5, 2.5, 3.5)

なお、『分類語彙表』に採録されていない特徴語と固有名詞は、

この分析の対象から除外した。

面種ごとに①～⑤の意味分野に属する特徴語の割合を求め、高頻度外来語全体の場合と比較したのが図7である。高頻度外来語全体では、各意味分野の割合は、①24.4%、②15.6%、③34.3%、④21.0%、⑤4.7%である。

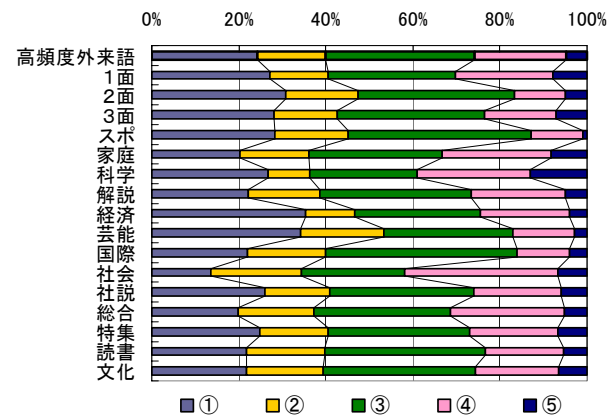


図7：面種ごとの各意味分野に属する特徴語の割合

以下、高頻度外来語全体における①～⑤の割合を基準としたときに、特にはっきりとした特徴を示した、スポーツ、家庭、科学、経済、芸能、国際、社会、総合の各面について、具体的な特徴語例を挙げながら考察する。

まずスポーツ面は、③の語が41.6%と際立って多い点が特徴的である。「ハンドボール」「ホッケー」「バドミントン」などの競技名、「バーディー」「ファウル」「シュート」などのスポーツ用語がここに含まれる。またもうひとつの特徴として、⑤の語が極端に少なく「リズム」の1語のみであることが挙げられる。

次に家庭面は、④の語が25.1%と比較的高い。「サラダ」「スープ」「チーズ」など食品関係の語、「スカート」「シャツ」「ネクタイ」など衣料関係の語、「ベランダ」「トイレ」「ベッド」など住宅関係の語が、それぞれ④の語として多数挙がっている。

科学面は、①の語が26.8%、③の語が24.7%、④の語が25.9%と均衡している点が特徴的である。①は「メカニズム」「データ」「サイクル」など、③は「エネ³」「プロジェクト」「プラント」など、④は「ロボット」「シャトル」「センサー」などである。また、科学面は、⑤の割合が13.0%と非常に高い点も特徴として挙げられる。「フロン」「オゾン」「プルトニウム」などの物質名や「ホルモン」「インフルエンザ」「アレルギー」といった生体に関する語が特徴語に多数含まれることの表れである。

経済面は①の語の割合が非常に高い(35.4%)という特徴を示した。「シェア」「ビッグバン」「レート」「シフト」など経済でよく使われる語のほかに「ドル」「マルク」「ポンド」といった、通貨単位を表す語が多数挙がっている。

芸能面もまた、①の語の割合が34.1%と高く特徴的である。「ジャンル」「ヒット」「シーン」「デビュー」などである。また、

² 宝くじの一種。当選の情報が新聞に定期的に掲載される。

³ 「省エネ」「エネ庁」など、「エネルギー」の略。

芸能面の特徴語には、②の語も 19.2%と比較的多い。ここには「ピアニスト」「アイドル」「アーティスト」「ヒロイン」といった芸能関係の職種や身分を表す語が多数含まれる。

国際面は、特徴語の大多数が固有名詞で作業から除外されたため、極端に少ない数の分析となったが、「クーデター」「ゲリラ」「デモ」など③の語の割合が際立って高い(44.0%)という特徴を示した。

社会面は、④の語が 35.1%と非常に多いことが特徴である。「ワゴン」「タンカー」「バイク」「トラック」といった乗り物や、「ナイフ」「ロープ」「テープ」「バッグ」のような用具類がここに含まれる。また、「スナック」「ホームレス」「クリニック」「ストーカー」など、②の語の割合が 20.9%と比較的高いことも社会面の特徴として挙げられる。

最後に総合面は、④の語の割合が 26.1%と比較的高い点の特徴的である。「ハイビジョン」「スタジオ」「テレビ」といったテレビに関する語のほか、「ラーメン」「ワイン」「コーヒー」といった食品関係の語が④に属する特徴語として多数挙げられた結果である。

6. まとめ

以上、特化係数により各面の特徴語を求め、それにしたがって、各面種ならびに高頻度外来語の考察を行った。

各面に占める特徴語の割合の比較からは、特徴語の異なり語数は総合面が最も多く芸能面が最も少ないこと、延べ語数はスポーツ面が最も多く総合面が最も少ないこと、国際面の特徴語の大部分を固有名詞が占めることなどを明らかにした。次いで特徴語の範囲の分析からは、固有名詞の特徴語となる範囲が一般語に比べ狭いこと、スポーツ面の特徴語に範囲1というスポーツ面に特化した語が多いことを明らかにした。最後に特徴語を意味分野によって整理することで、スポーツ、家庭、科学、経済、芸能、国際、社会、総合の各面でそれぞれ特徴的な性質を見出すことができた。

本研究では、特化係数が標準値1を超える語をすべて各面の特徴語として同列に扱ったが、特徴語の特化係数の値は、1をわずかに超えるものから 20、30 といった極端に大きい値までさまざまである。面種ごとの特化係数の分布を確認した上で、特徴語とする値を面ごとに1よりも大きい値に設定すれば、分析結果から得られる面種や語の特徴は、よりはっきりとしたものになると考えられる。今後は、それを試みることで、特徴語となる高頻度外来語のより具体的な考察を行いたい。

また、本研究では高頻度外来語を通じて新聞の面種とその語彙の分析を行ったが、新聞の語彙の性質を明らかにするという目的のためには、外来語以外の語種に分析の対象を広げることが不可欠である。それについても、今後の課題としたい。

参考文献：

- 国立国語研究所(1995・1999)：
『テレビ放送の語彙調査 I・III』秀英出版
山崎誠 (2005)：
『現代雑誌の語彙調査』第 28 回語彙・辞書研究会

参考：特徴語例（特化係数の高い順）

- [1面] ダウ ナスダック クローズアップ トウパク・アマル ヘリポート ノンバンク ハイテク トウパク ナトリウム レーダー …
- [2面] リベラル ヘリポート プロジェクトチーム ガイドライン ゴラン スタンス ポスト ビジョン ヒアリング ジレンマ …
- [3面] ビルマ ベア ブルサーマル クロウン フロン プルトニウム チェルノブイリ ヘリポート キロワット ナトリウム …
- [スポーツ] スライダー コールド サッカーリーグ ダート センタリング バックス スクイズ ミキハウス ボギー …
- [家庭] ボシェット リフォーム レシピ サラダ ビタミン シニア アレルギー トマト スープ ホルモン ヘルパー カルシウム …
- [科学] マウス プレート フロン ロボット レーザー シャトル オゾン ステーション リサイクル エネ メカニズム …
- [解説] コラムニスト メールアドレス マナー ニュース インフォメーション パート リポート ワープロ ノート モラル ペン…
- [経済] ソフトバンク ビジネス ルノー トレンド メモリー ドコモ マツダ デリバティブ アサヒ アナリスト シェア …
- [芸能] ソプラノ バッハ ポップス ベートーベン モーツァルト チェロ サウンド アクション アルバム ミュージカル ピアノ…
- [国際] アチェ ピノチェト バラク ワヒド メガワティ ハンナラ ハタミ スーチー カシミール マケドニア シャロン デイリ …
- [社会] セット・ボックス セット・ストレート ナンバーズ ドリーム ボックス マグニチュード スナック ワゴン ナイフ オウム …
- [社説] マニフェスト パート モラル レジ ビジョン マナー リーダーシップ グローバル プルトニウム ノー ロッキード …
- [総合] ニュースランキング デイキャッチ クロスロード ダッシュボード フジ ワイド スペシャル ファイル サロン テレ …
- [特集] ハウジング・ミニ コンテンツ バリアフリー ハウス リフォーム デザイナー コレステロール ユース コレクション …
- [読書] インフォメーション ベストセラー ミステリー コラム ノンフィクション エッセー ページ イギリス ダイヤモンド …
- [文化] ギャラリー アート モチーフ ジャンル ミステリー シンポジウム シンポ ダイナミック ノンフィクション バッハ …