日本語流行歌中の外来語に関する調査

吉田桃子
東京外国語大学大学院　地域文化研究科

望月源
東京外国語大学　外国語学部

1　はじめに

本研究は流行歌の歌詞を手がかりにして，日本語における外来語の役割を，その量と使われ方の変化から明らかにすることを目的としている．

一般に外来語は，大きく分けて1．これまでになかった事物を言い表す，2．洗練された雰囲気を作りだす，といった目的で用いられることが多い．外来語が多く使われる分野として，いわゆる「お役所文書」や「歌詞」が挙げられる．前者においては情報伝達が第一の目的なので，外来語の多用が分かりにくさにつながるとしてしばしば問題になる．一方後者においては，内容は基本的に自由である．語や表現の選択に制約はないが，多くの人に受け入れられたものが流行歌となる．このことから日本語流行歌に使われる外来語の量と内容は日本人が外来語に対して持つイメージや理想を極端な形で反映していると考えられる．したがって流行歌歌詞の外来語使用を調査することは，外来語一般のイメージを明らかにする一つの手段となりうる．

従来から流行歌を対象にした語彙調査は多いが，外来語に重点を置いたものは少ない．外来語を調査対象としたものでも，他の語種を考慮していなかったり，調査対象期間が限られていたりするものがほとんどである[1][2][3]．本研究では歌詞全体に占める外来語の実質的割合を知るため，歌詞のすべての自立語を数え上げの対象とする．また外来語使用の量的推移をよりはっきりした形で観察するため，明治元年から平成14年上半期までの135年という比較的幅のある期間を調査対象とする．

本調査では，流行歌歌詞における外来語の量的変遷，外来語の品詞，外来語が多く使われる意味分野の変遷，具体的な高頻度外来語について報告する．

2　単語の認定と語種の定義

2.1　単語の認定基準

日本語における「単語」の定義は研究者によって様々である．そのため語彙調査では「単語」をどのような単位で設定するかが常に問題となる．国立国語研究所の提案する調査単位には「文節（単語）相当の長い単位系」のものと「形態素相当の短い単位系」のものがあり，先行の語彙調査でそれぞれ採用，応用されている．

本調査では「長い単位」を採用する．「長い単位」は文節を認定して，そこから自立語を取り出すという手順で決定されるので「短い単位」よりも認定のばらつきが少ないからである．

「単語」の同語異語判定も問題になるが，調査全体に与える影響は「単語」自体の定義よりも小さい．本調査では文字テキストを基準にした処理が原則なので，語のまとめあげは表記によって行う．表記（ひらがなと漢字など）が違っているだけで，同語だと判定したものは同一の見出しにまとめる．ただし，今回の調査対象の中で2種類以上の漢字表記があるものはまとめない．例えば「青い」と「蒼い」は別々の見出し語とする．さらに「あおい」があった場合はまた別の見出し語とする．「青い」と「あおい」だけがあり，他に「あおい」と読む漢字表記がない場合にだけ「青い」と「あおい」をまとめて同語とする．

2.2　語種の定義

語の原籍に基づいて語彙を分けていく分類法を語種分類という．一般的に日本語の語種分類は「和語」，「漢語」，「外来語」の3分法をとる[4]が，本研究では「外来語」との対比において「和語」と「漢語」を区別せずに「和語・漢語」とする．また，歌詞には日本語以外の言語種も使われているため，これらをまとめて「外国語」として扱う．さらに「外来語」と「和語・漢語」によって構成される複合語を「外来混種語」とする．以上，本調査の語種・言語種分類は「外来語」「外国語」「外来混種語」「和語・漢語」の4種類である．なお「外来語」，「外国語」の定義は文献[5]に準じ，以下のように定義する．

外来語：仮名かローマ字表記の外国出自の単語
外国語：アルファベット表記の外国出自の単語

3　調査方法

調査の具体的な手続きは次のとおりである．歌詞全体を単語に区切り，語種，原語，品詞，意味コードなどの情報をつけた語彙表を作成する．付加された情報

に基づいて単語の数え上げを行い，得られた数値を分析する．なお単語の切り出しのために係り受け解析器「南瓜 windows 版[1]」を用いる[6].

3.1 歌詞の収集

調査対象期間は明治元（1868）年から平成 14（2002）年上半期までの 135 年とし，ほぼ 10 年ごとに 14 の年代に区切る．調査対象曲は，1868 年から 1967 年までは歌本『日本のうた』[7]から 1 年代につき 50 曲を上限に無作為抽出，1968 年以降は『オリコン年鑑』[8]の年間売上ランキングから 1 年あたり上位 5 曲を選ぶ.

このようにして得られる調査対象曲は合計 584 曲であり，年代区分と対象曲数は表 1 のとおりである.

表 1 年代区分と対象曲数

	期間			年数	曲数
1.	1868（明治元）	～	1874（明治 7）	7	8
2.	1875（明治 8）	～	1884（明治 17）	10	14
3.	1885（明治 18）	～	1894（明治 27）	10	28
4.	1895（明治 28）	～	1904（明治 37）	10	44
5.	1905（明治 38）	～	1914（大正 3）	10	50
6.	1915（大正 4）	～	1924（大正 13）	10	50
7.	1925（大正 14）	～	1934（昭和 9）	10	50
8.	1935（昭和 10）	～	1944（昭和 19）	10	50
9.	1945（昭和 20）	～	1954（昭和 29）	10	50
10.	1955（昭和 30）	～	1964（昭和 39）	10	50
11.	1965（昭和 40）	～	1974（昭和 49）	10	50
12.	1975（昭和 50）	～	1984（昭和 59）	10	50
13.	1985（昭和 60）	～	1994（平成 6）	10	50
14.	1995（平成 7）	～	2002（平成 14）	8	40

3.2 前処理

流行歌の歌詞は通常の書き言葉とは異なり，以下のような話し言葉に近い性格を持っている.

1. 意味のない繰り返しがある.
2. 倒置が頻繁に起こる.
3. 助詞・助動詞の脱落が多い.
4. 文が完結しない，あるいは他の語句の割り込みによって中断することがある.

これらの特徴を持つテキストは，そのままでは「南瓜」による解析がうまくいかない．解析精度を上げるために歌詞テキストに対し以下のような前処理を行う.

- 旧仮名遣いを現代仮名遣いに（「言ふ」→「言う」），旧字体を新字体に，標準的でない，あるいは「茶筅」で認識されない表記を適当な表記に（「たぐい」→「類」，「ワルいヤツ」→「悪い奴」）変える.
- 助詞の脱落により文節区切りが明確でない箇所や語が並列されている箇所を区切る（「紅い帯締め」→「紅い帯（を），締め」）.

3.3 文節の取り出し

「茶筅」と「南瓜」を使って歌詞を文節単位に区切り，形態素に品詞情報をつける．「南瓜」でうまく区切れなかった文節は，付加された品詞情報を利用して，プログラムにより修正する．その際適用した規則は主に以下のようなものである.

一つの文節中に以下のような品詞の連続がある場合はその間を区切り，別々の文節にする.

- 「名詞」と「動詞」（「唄/ひびく」）．ただし「名詞-サ変接続」と「名詞-形動語幹」に接続する「スル」は切らない（「仕事する」「寝坊する」）.
- 「助詞，助動詞」と「名詞，未知語，動詞」（「大将たる/者」「好きに/なる」）.
- 「名詞-接尾（助数詞は除く）」と「名詞（接尾と非自立は除く），動詞」（「見ずや/夕暮れ」「見合わすネ/顔と」）（「十八・年・目・で」は切らない）.
- 「名詞」と「連体詞」，「連体詞」と「名詞」（「拙者/この/町に」）.
- 「名詞」と「外国語（＝アルファベットだけで構成されている未知語）」（「乾杯/BABY」）.
- 「形容詞」と「動詞」（「早く/来い」）.

文節が必要以上に区切れてしまっている以下のような箇所は，直前の文節とつなげる.

- 「助詞，助動詞」で始まる文節.
- 「形式動詞（「…という」の「いう」など）」で始まる文節.

上記の規則を適用してもうまく区切れなかった部分をチェックし，手作業で修正する．アルファベット語（本調査における外国語）は「未知語」に区分されるので，この段階で文脈上の品詞を付加する.

3.4 「長い単位」の語の取り出し

一つの文節から一つの自立成分またはそれに準ずるものを取り出す．複合語は形態素に付加された情報を手がかりに認定し，複合した形で取り出す．複合語と規定するのは以下のような品詞の連続である.

- 「名詞」と「名詞，未知語」.
- 「動詞」と「動詞，名詞，形容詞-非自立」（「押し

1 「南瓜」用の形態素解析器には「茶筅 2.1 for windows」を用いている.

+出す」「くわえ+煙草」「住み+よい」).

- 「名詞，未知語，副詞」と「スル」.

3.5 語彙表の作成

3.4 節で得られた「長い単位」の単語リストから，2.1 節で述べた基準により同じと認められるものは同一の見出し語の元にまとめて，見出し語リストを作成する．重複する出典情報（どの歌に出現するか）をまとめて，出現回数をつける．外来語，外来混種語，外国語と判断したものに印をつける．外来語に分類したものには，分類語彙表（増補版）[9]の意味コードをつける．最終的に得られるデータの内容と形式は以下のとおりである．

通し番号；語種；見出し語；品詞；出現形；出典情報；意味コード

4 結果と考察

調査対象全体から，異なりで 9,397 語，述べで 48,231 語が得られた．語種の内訳を表 2 に示す．

表 2 年代を通じての語種別語数

	異なり	延べ
和語・漢語	8,446	43,741
外来語	486	1,141
外来混種語	106	160
外国語	359	3,189

和語・漢語以外の語種が年代を通じて歌詞全体に占める割合は，異なり，延べともに 10%程度である．それぞれの語種が歌詞の異なり語数に占める割合を年代別に示すと図 1 のようになる．

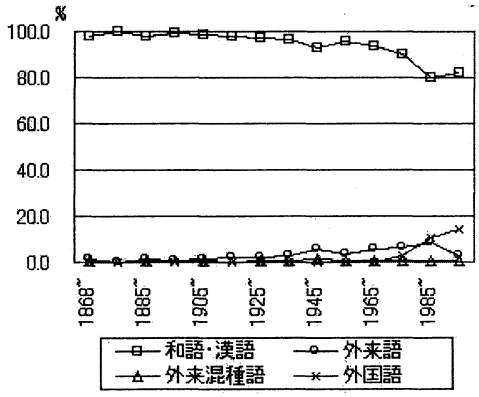

図 1 異なり語数に占める語種の割合

全年代を通じて歌詞の大部分を占めているのは和語・漢語である．外来語の割合は歌詞全体から見れば多くはない．1945 期の増加と 1995 期の減少が多少目立つ程度である．外来語に代わって 1975 期から増加してくるのが外国語である．外来語が減って外国語が増えているのは，外来語が担っていた装飾的機能（1 節で述べた外来語使用の目的 2）が外国語に移行したからだと考えられる．つまり外来語の利用価値が下がって外国語の利用価値が相対的に上がったということである．1995 期に和語・漢語が増える兆候が見られるのは，和語・漢語の意味伝達機能が見直されたためだと考えられる．このように，それぞれの語種はその機能や利点によって使い分けられていると推測できる．

外来語の品詞は名詞がほとんどである．一方，外国語の品詞を使用量の安定する 1975 期以降について見ると，名詞も多いが，その他の品詞も多いことが分かる（図 2）．これは外来語が単語単位で取り入れられることが多いのに対し，外国語が文章単位で取り入れられることが多いことの表われである．

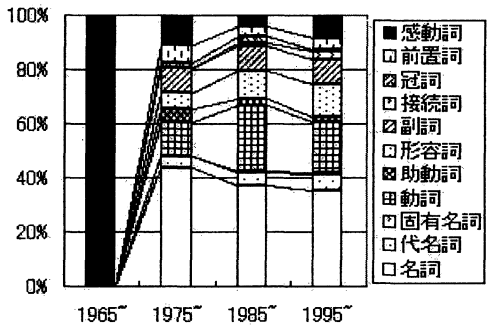

図 2 外国語の品詞(異なり)

外来語がほとんど名詞であることは上に述べた．流行歌に使われる外来語名詞の意味を，分類語彙表による分類に従って意味コードの 2 桁目まで分類した結果を図 3 に示す．意味コードの内容は表 3 のとおりである．

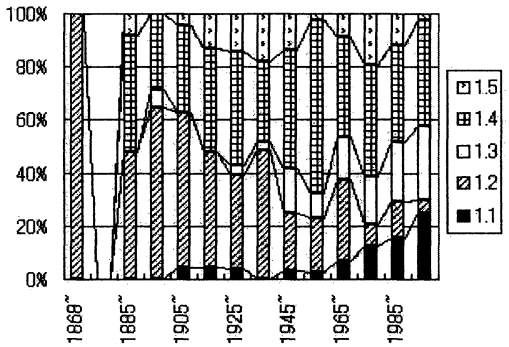

図 3 外来語(異なり)体の類下位区分の割合

表 3 分類語彙表の意味コード

1.	体の類(名詞類)
1.1	抽象的関係
1.2	人間活動の主体
1.3	人間活動 精神および行為
1.4	生産物および用具
1.5	自然物および自然現象

どの年代の外来語にも多いのは,具体物を示す名詞(1.4 の分類)である.ただし 1945 期以降は抽象的な語(1.1, 1.3 の分類)の割合が増えている.その理由は外来語一般に抽象語が多くなってきたことと,流行歌における外来語の使い方が変わってきたことが考えられる.雰囲気を演出する小道具として使われる外来語の他に,雰囲気や感情を直接表現する外来語が使われるようになってきたのである.

具体的な高頻度外来語を見ても上に述べたような傾向が分かる.年代別の高頻度外来語を表 4 に示す.高頻度語は通常,延べ語数から計算した使用率によって示されるが,歌詞特有の繰り返しの影響を避けるため,一つの年代について複数曲に登場する外来語を示している.なお 1868〜1885 期の 3 つの年代は複数曲に出現する外来語が一つもない.

表 4 年代別,複数曲に出現する外来語

1895 期	ラッパ, トンネル
1905 期	バイロン
1915 期	シナ
1925 期	パリ, ポプラ, アカシア
1935 期	タバコ, シナ, ランタン
1945 期	グラス, ホテル, ハンカチ, パイプ, ブルース, ランプ, テープ
1955 期	ギター, ビル, ネオン
1965 期	ドア, タバコ, ブルー, レース(飾り)
1975 期	キス, タバコ, ボトル
1985 期	キス, ガラス, コイン, リズム, ベッド, ベル, ラジオ, ダイヤモンド, スキャンダル, エクスタシー, オレンジ, キャンドル・ライト, ターン, バラード
1995 期	ガラス, ドア, キス, ベル, メロディ, リスク, コロン, ストーリー

1935 期以前は地名・人名の固有名詞や具体名詞がほとんどである.1945 期以降は固有名詞が見当たらなくなり,一般名詞がほとんどになる.その中には具体名詞とともに抽象名詞が現われている.1935 期以前は外来語の数自体が少なかっただけなく,和語・漢語で置き換えにくい,つまり必要のため(1 節で述べた外来語使用の目的 1)に使われる外来語が多かったということである.

歌詞に使われる外国語はほぼすべてが英語であった.語の種類という点からは,外来語より外国語に定型化の傾向がみられる.どの年代においても使用頻度が高い外国語は「I」と「you」である.これは近年の流行歌で描かれるのが基本的に「私」と「あなた」の世界であるということと,日本人の考える外国語(=英語)の構文的,語彙的バリエーションがまだ少ないということを意味していると考えられる.

5 おわりに

流行歌の歌詞を「単語」に区切り,語種,言語種別に分類,それぞれの使用量を調べた.外来語の割合はそれほど多くなく,外国語の割合が近年になって急増していることが分かった.また,語種はその機能によって使い分けられているということが分かった.

今後,調査対象曲数を増やして和語・漢語も含めた分析を行うことによって,語種の使い分けの原則,ひいては使用者の意識や語と語種に対するイメージを明らかにすることができると考えられる.

6 参考文献

[1] 堀江真喜雄 (1966)「『ネオン』と『ギター』外来語使用から見た歌詞の一側面」『言語生活』178, pp.70-74 筑摩書房

[2] 原忠彦 (1977)「歌謡曲における漢字音語と外来語」『アジア・アフリカ言語文化研究所通信』31, pp.5-9 東京外国語大学アジア・アフリカ言語文化研究所

[3] 米田武 (1980)「流行歌の中の外国語」『文研月報』30-7, pp.2-38 日本放送協会

[4] 田中章夫 (1978)『国語語彙論』 明治書院

[5] 伊藤雅光 (2001)「ポップス系流行歌の語彙調査における外来語と外国語の判定基準」『計量国語学』23-2, pp.110-130 計量国語学会

[6] 工藤拓・松本裕治 (2002)「チャンキングの段階適用による日本語係り受け解析」『情報処理学会論文誌』43-6, pp. 1834-1842

[7] 野ばら社編集部・椎葉京一 (1998-2001)『日本のうた』第 1 集〜第 7 集 野ばら社

[8] 小池聰行 編 (1968-2001)『オリコン年鑑』 オリコン

[9] 国立国語研究所・中野洋 (1996)『「分類語彙表」形式による語彙分類表(増補版)』第 1・2 分冊 国立国語研究所