

平成21年5月25日現在

研究種目：基盤研究（C）
 研究期間：2005～2008
 課題番号：17520271
 研究課題名（和文） 身体部位語の連語関係記述を基にした「学習」和独辞典のための基礎的研究
 研究課題名（英文） Research on an ideal learner's Japanese-German Dictionary based on the description of idioms and collocations
 研究代表者
 恒川 元行（TSUNEKAWA MOTUYUKI）
 九州大学・大学院言語文化研究院・教授
 研究者番号：70197747

研究成果の概要：

本研究の成果は、以下の4点である：①ドイツ語辞書記述（独和、和独）の土台となる語彙データ入手手順の確立および語彙調査支援ソフトの開発、②5つのテキストを対象にした語彙調査の実施、③独自の語彙調査結果を検証・補足する目的での大規模ドイツ語コーパス（ベルリン・ブランデンブルク科学アカデミーのDWDSコーパス等）の利用法調査、④平成21年度科研費基盤研究（C）「ドイツ語テキスト及び文における語彙出現予測分析とその和独辞典・教材への応用」（研究代表者：三重大学人文学部・井口靖教授）への展開。

交付額

（金額単位：円）

	直接経費	間接経費	合計
2005年度	700,000	0	700,000
2006年度	500,000	0	500,000
2007年度	500,000	150,000	650,000
2008年度	500,000	150,000	650,000
年度			
総計	2,200,000	300,000	2,500,000

研究分野：人文学

科研費の分科・細目：言語学・言語学

キーワード：学習辞典、和独辞典、語彙調査、コーパス

1. 研究開始当初の背景

（1）以前の科学研究費補助研究「身体部位・身ぶりに関わるドイツ語連語関係記述のための基礎的研究」（研究代表者：恒川元行、基盤研究（C）（2）、課題番号：13610629、2001～2004年度）では、身体部位語（Auge, Ohr,

Naseなど）をキーワードとして、その特徴的な慣用語法を捕捉することはもちろん、むしろより一般的な連語関係に注目し、当該語が極めてありふれた使われ方においてどのような連語関係を示すか、という問題提起を中核に据えた。

(2) ドイツ語辞書(独辞典)においても、近年では連語関係に配慮した辞典が編纂され始めている(たとえば Langenscheidts Großwörterbuch Deutsch als Fremdsprache における、独立した「連語関係記述欄」の新設)。しかし、従来、一般的連語関係研究は主に英語圏で盛んであり(『BBI 英和連語活用辞典』1993 など)、ドイツ語圏では慣用語法研究(H. Schemann: Deutsche Idiomatik. 1993 など)に比べそれほど注目を集めてこなかった。上記科研費補助研究においてドイツ語身体部位語の一般的連語関係を研究対象としたのは、このような動機からであった。

(3) 日本でのドイツ語初学者を対象とした学習和独辞典では、一般連語関係を記述した独辞典が出版され始めたことを受け、慣用語法研究、連語関係研究の成果が十分でないとしても徐々に記述に反映され始めている。これに対し、編集者・出版社にとって経済的に負担が重く、改定の機会がより乏しい和独辞典の場合、一般連語関係に配慮した辞典、さらにはこの観点からの初学者用学習辞典の編纂は、当時も現在も行われていない。

2. 研究の目的

(1) 本研究は、当初、上記科研費補助研究の成果であるドイツ語の身体部位語(Auge, Ohr, Nase など)をキーワードとする慣用語、連語関係の記述を手掛かりとして、和独辞典の編纂を理論的・方法論的に考察すること、また身体部位語に特化した小規模な和独辞典を作成することを目的としていた。

(2) しかし、実際の作業の中で、身体部位語を出発点とする考察では一般性に欠け、たとえば初学者が必要とする和独辞典にはどのような見出し語が必要か、また学習辞典の

基本として重要語と一般語で記述の質的区別を行う目的でも、どのような語が重要語になりうるか、などの問題提起に答えることができないことが明らかになった。そのため、2007年度より、言語データに裏付けられた見出し語の抽出と辞書記述に資することを目的として、コンピュータを用いた語彙調査方法を開発・導入し、ドイツ語(教育)基本語彙、特定分野語彙等、語彙の質的違いを明らかにしようとする、より一般的な語彙調査研究の方向へと展開した。

3. 研究の方法

(1) コンピュータを利用した語彙調査

ドイツ語(教育)基本語彙、特定分野語彙等、語彙の質的違いを明らかにする目的から、以下の異なる内容・分野の5つのテキストを対象として語彙調査を実施した:

① 『ドイツ人が日本人によく聞く100の質問』(三修社、2001)

*日本の衣食住習慣等をドイツ人に説明する内容のテキスト

② Andrea Raab u. Toshiko Ishii: Neugigkeiten aus Deutschland 07. (朝日出版社、2008)

*2007年のドイツでの出来事を紹介する時事的テキスト(スポーツ、文化、経済、社会、国際政治各2篇ずつ計10篇)

③ Wikipedia の記述(2. Sept. 2008)

„Globale Erwärmung“

*「地球温暖化」に関する百科辞典的解説テキスト

④ „Hunde kennen kein >richtig< oder >falsch< “. in: Jean Donaldson: Hunde sind anders. Stuttgart, 2000. 4. Kap., S. 129-170

*犬の行動の特徴に関するテキスト

⑤ Resolution 1 des Weltgipfels für soziale Entwicklung. Anlage I:

Kopenhagener Erklärung über soziale Entwicklung, 1995.

*1995年の国連「社会開発に関するコペンハーゲン宣言」第1部

(2) 自動化の開発・導入

語彙調査を統一的基準で実施し、頻度集計や使用率計算を自動化して省力化を図るため、また複雑な調査過程での漏れや間違いを排除するため、協力を得て以下のエクセルのマクロを作成した：

- ①「要注意語を語尾でチェック」
- ②「要注意語チェック」
- ③「重複語のチェック」
- ④「章別頻度の計算」
- ⑤「章数と合計頻度と平均使用率の計算」
- ⑥「重複語の整理」
- ⑥ 全体に合わせて各章の見出し語を整列」

(3) 研究会等

各種の研究会において発表を行ったほか、学会等を利用して和独辞典、語彙調査などに関する意見交換、議論の機会を設けるよう積極的に努めた。この意見交換の成果は、本研究に直接反映されただけでなく、平成21年度科研費基盤研究(C)「ドイツ語テキスト及び文における語彙出現予測分析とその和独辞典・教材への応用」(研究代表者：三重大学人文学部・井口靖教授)への契機の一つとなった。

(4) 大規模ドイツ語コーパスの利用研究

パーソナルコンピュータ規模の語彙調査だけでなく、その調査結果を検証・補足する目的から、また和独辞典研究の今後の展開を念頭に、大規模ドイツ語コーパスの一つとして DWDS コーパス(ベルリン・ブランデンブルク科学アカデミー)、また Institut für

Deutsche Sprache(ドイツ語研究所)コーパスの利用方法を調査した。

4. 研究成果

本研究の成果は、①ドイツ語辞書記述(独和、和独)の土台となる語彙データ入手手順の確立および語彙調査支援ソフトの開発(公開準備中)、②5つのテキストを対象にした語彙調査の実施(発表準備中)、③独自の語彙調査結果を検証・補足する目的での大規模ドイツ語コーパス(ベルリン・ブランデンブルク科学アカデミーの DWDS コーパス等)の利用法調査(今道・恒川 2009)、④平成21年度科研費基盤研究(C)「ドイツ語テキスト及び文における語彙出現予測分析とその和独辞典・教材への応用」(研究代表者：三重大学人文学部・井口靖教授)への展開の4点である。

以下、本節では、成果の具体例として、②の語彙調査結果を略述する。

(1) 「異なり語数」と「延べ語数」

語彙調査の対象とした5つのテキスト(「3. 研究の方法(1) コンピュータを利用した語彙調査」を参照)の「異なり語数」および「延べ語数」は、表1のとおりである。

(表1)

	異なり語数	延べ語数
①100Fragen	2278	10858
②Neuigkeiten	1438	3645
③Erwärmung	1671	6588
④Hunde	2017	12410
⑤Entwicklung	1569	9588
全体	6347	43089

5つのテキストのうち、①、④はテキスト規模(延べ語数)も大きくテキスト内で使用されている語彙も豊富であること(異なり語

数)、⑤はテキスト規模に比べ使用語彙が相対的に乏しいことがわかる。

(2) 「深さ」と「広さ」による調査語彙の分布

基本語彙は単に頻度(使用率)の高さだけで決めることができない。そのため、樺島・吉田(1971)、林(1975)にならい、使用頻度(以下「深さ」)のほか、第2の指標として当該語が調査テキストのうちいくつに共通して出現するか(以下「広さ」という観点を加え、語彙分布を調査した。

以下の表2の縦軸の5~1は「深さ」を示し、今回の調査では使用率4.00%以上を「5」、2.00%以上4.00%未満を「4」、1.00%以上2.00%未満を「3」、0.50%以上1.00%未満を「2」、0.50%未満を「1」とした。

また、表2の横軸の5~1は「広さ」、すなわちたとえば「5」は当該語が5つの調査テキスト全部に共通して出現していることを示す(以下同じ)。

(表2)

	5	4	3	2	1	計
5	28	2	2	4	49	85
4	21	9	13	13	141	197
3	43	42	67	97	825	1074
2	45	93	165	279	932	1514
1	0	37	112	418	2910	3477
計	137	183	359	811	4857	6347

文献:

樺島忠夫・吉田弥寿夫「留学生教育のための基本語彙表」大阪外国語大学「日本語・日本文化」第2号、1971.

林四郎「第二章 基本語彙はきめられるか／第四節 新聞基幹語彙から教育基本語彙をかいまみる」『新・日本語講座1 現代日本

語の単語と文字』汐文社、1975.

(3) 各グループの語彙

以下、例として、具体的に「深さ5/広さ5」(以下5-5)、「深さ5/広さ1」(以下5-1)、「深さ3/広さ3」(以下3-3)グループの語彙を示し、特徴を述べる。

① 5-5グループの語彙(28語)

auch, nicht (副詞2語=7.1%);
 geben, sein (動詞2語=7.1%);
 all, d, dies, ein(冠詞類);als, dass, oder,
 und(接続詞); an, auf, bei, für, in, mit,
 von, zu(前置詞); es, ihr, sich, sie(代
 名詞); haben, können, werden(助動詞);
 wie(疑問詞)(その他の品詞24語=85.7%)

このグループの特徴は、冠詞類、代名詞、前置詞、接続詞、助動詞、数詞、疑問詞など(以下「その他の品詞」)の割合が高い(85.7%)ことである。

(表3)

	5	4	3	2	1
5	85.7%	50.0%	0	0	0
4	57.1%	22.2%	23.1%	7.7%	0.7%
3	32.6%	23.8%	14.9%	12.4%	3.4%
2	13.3%	14.0%	12.1%	8.2%	5.4%
1	0	2.7%	2.7%	6.2%	4.7%

高頻度語には機能語と呼ばれるこのような「その他の品詞」が多く、これらの語はまた同時に、幅広くいずれのテキストにも共通して見られる語彙である。表3は、このことを端的に示している。

これに対し、本グループには名詞は含まれず、2例ずつ見られる形容詞・副詞、動詞も

典型的な形容詞・副詞、動詞ではなく、機能的な内容の語と見ることができる (auch, nicht; sein)。

② 5-1 グループの語彙 (49 語)

Babyboom, Baseball, Beschäftigte (der), Bundeskanzlerin, Doping, Erwärmung, Fahrer, Fisch, Gaststätte, Gemüse, Gericht, Geschenk, Handball, Japaner, Kämpfer, Kinderbetreuung, Kommune, Konzern, Krippenplatz, Radsport, Rauchverbot, Sponsor, Team, Teezeremonie, Urlaub, Weltmeisterschaft, Werk (一般名詞 27 語=55.1%) ;
Airbus, Bausch, Brand, EADS, Heiligendamm, Japan, Löscher, Pina, Power eight, Siemens, T-Mobile, von der Leyen, ZDF (地名・人名・企業名等 13 語=26.5%) ;
Fuji, Ikebana, Kabuki, Kanji, Kimono, Noh, Shinkansen (日本語 7 語=14.3%) ;
rauchfrei (形容詞 1 語=2.0%) ;
bellen (動詞 1 語=2.0%)

本グループの語彙は、5-5 グループと同じく高頻度ではあるが、5 つのテキストのひとつにしか見られない「広さ」の狭い語である。このグループには名詞類 (一般名詞、地名・人名・企業名等、日本語) が 95.9% と圧倒的に多く含まれており、その特徴となっている。また、この名詞類自体にも Airbus、Siemens などの地名・人名・企業名等、Ikebana、Kabuki などの日本語語彙が多数 (42.6%) 含まれ、ひとつの特徴をなしている (特定分野語彙)。

5-5 グループと 5-1 グループのこのような質的相違は、語彙調査の尺度として単に使用率 (頻度) を用いるだけでは不十分であり、少なくとも「広さ」のような第 2 の指標が不

可欠であることを示している。

③ 3-3 グループの語彙 (67 語)

Arbeitsplatz, Atmosphäre, Ding, Fernsehen, Geschäft, Gesellschaft, Herbst, Jugendliche (der), Kampf, Kleinkind, Klima, Leute, Mal, Mann, Meter, Minute, Nation, Politik, Quelle, Schule, Sommer, Spiel, System, Tag, Thema, Zeitung, Ziel (名詞 27 語=40.3%) ;
bedeutend, best, deshalb, deutlich, einfach, einmal, europäisch, fast, heiß, klassisch, kurz, nämlich, öffentlich, rechtlich, richtig, stark, vereint, weltweit (形容詞・副詞 18 語=26.9%) ;
bekommen, bewegen, dauern, leben, lernen, messen, nehmen, schenken, sprechen, stärken, tragen, unterstützen (動詞 12 語=17.9%) ;
acht, dritt, fünfzehn, null, sechs, vierundzwanzig (数詞) ; etwas (代名詞) ; pro (前置詞) ; unser (冠詞類) ; warum (疑問詞) (その他の品詞 10 語=14.9%)

本 3-3 グループは、「深さ」「広さ」ともに中央に位置する語彙グループである。ここには名詞、形容詞・副詞、動詞、「その他の品詞」のいずれもが含まれているが、「その他の品詞」は相対的に少ない。名詞、形容詞・副詞、動詞には、それぞれ Fernsehen, Geschäft, Gesellschaft ; einfach, einmal, europäisch ; dauern, leben, lernen など、きわめて基本的と考えられる語が含まれている。

(4) まとめ

「深さ」と「広さ」の 2 つの指標に基づく語彙調査は、以上のように語彙分布の質的な

相違を明らかにするための有効な手段であることがわかった。本研究をとおしてパーソナルコンピュータを用いた語彙調査の手法を確立できただけでなく、実際、語彙分布に関する興味深い事実を明らかにした。

しかし、この成果を和独辞典、学習辞典の編纂に生かすためには、本研究の基礎の上に調査手法の検証と洗練、さらなる語彙調査や大規模ドイツ語コーパスのデータ利用などの課題が残されている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計3件)

- ① 今道晴彦・恒川元行、DWDS コーパスの概要と利用法 (研究報告)、九州大学大学院言語文化研究院『言語科学』、44、2009、147-166、査読なし
- ② 恒川元行、小規模な語彙調査に基づく基本語彙抽出の試み、九州大学言語文化研究院『言語科学』、43、2008、107-117、査読なし
- ③ 恒川元行、wischen/abwischen の用法と共起成分 (研究報告)、九州大学言語文化研究院『言語科学』、42、2007、77-90、査読なし

[その他]

○研究会発表 (計3件)

- ① 恒川元行、コンピュータを用いた語彙調査のデータ処理の自動化について、和独研究会、2008年9月16日、三重大学
- ② 恒川元行、小規模な語彙調査に基づく基本語彙抽出の試み、東京外国語大学ドイツ語学研究会、2008年2月8日、東京外国語大学

- ③ 恒川元行、和独辞典の採録語の選択について、和独研究会、2007年9月16日、三重大学

○語彙調査用ソフトの開発 (計7件)

語彙調査の省力化および信頼性向上のためのエクセルマクロ:

- ① 「要注意語を語尾でチェック」
- ② 「要注意語チェック」
- ③ 「重複語のチェック」
- ④ 「章別頻度の計算」
- ⑤ 「章数と合計頻度と平均使用率の計算」
- ⑥ 「重複語の整理」
- ⑦ 「全体に合わせて各章の見出し語を整理」

6. 研究組織

(1) 研究代表者

恒川 元行 (TSUNEKAWA MOTOYUKI)

九州大学・大学院言語文化研究院・教授

研究者番号: 70197747