

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 5 月 15 日現在

機関番号：11301

研究種目：基盤研究（C）

研究期間：2009～2011

課題番号：21520393

研究課題名（和文）コーパス利用のコロケーション記述の理論と方法に関する通言語的研究

研究課題名（英文）A Cross-linguistic Study on a Theory and Methods of Prescribing Collocations on the Basis of Corpora

研究代表者

後藤 斉（GOTOO HITOSI）

東北大学・大学院文学研究科・教授

研究者番号：90162156

研究成果の概要（和文）：コーパスからのコロケーション情報の抽出の手法および言語教育上で有意義な情報の整理について、理論的な考察とケーススタディを行った。コロケーション類似現象は統語的に密接な関係にある内容語同士の場合を典型としつつも、その周辺部においては、それぞれの言語における語の形態論的性質に依存して、多様な広がりを見せていて、コーパスからのコロケーションの抽出にあたっては、このことを十分に考慮すべきであるとの結論にいたった。

研究成果の概要（英文）：Theoretical considerations and case studies have been conducted on methods of retrieving collocation information from corpora and ways of arranging and presenting information in view of language teaching. Collocation-like phenomena are typically found between syntactically related content words, but widely varied in marginal areas depending of morphological statuses of words in the language in question. In retrieving collocations from corpora, this fact should be fully considered.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,200,000	360,000	1,560,000
2010年度	500,000	150,000	650,000
2011年度	300,000	90,000	390,000
年度			
年度			
総計	2,000,000	600,000	2,600,000

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：言語学、辞書論、コロケーション

1. 研究開始当初の背景

研究代表者は、早くから日本語コーパス言語学を構想し、その成立に向けて試行的な研究を行ってきた。また、財団法人日本エスペラント学会の委託を受けてエスペラント日本語辞典編集委員会に編集副主幹として加わり、2006年の『エスペラント日本語辞典』（日本エスペラント学会）の刊行に多大の貢献をした。これにあたっては、研究代表者が独自に構築したテキストデータベース

が大いに活用され、従来は記述の不十分であったコロケーションに関してもその成果を生かした記述をすることができた。この辞書編集における経験に基づく知見は、エスペラント以外の言語の研究にも生かす価値のあることであると考えられる。

英語コーパス言語学はその研究の展開のなかで理論的な知見と実践的な経験とを蓄積してきており、日本語など他の言語の研究におけるコーパス利用にあたっては参考

なる点が数多い。エスペラントは計画言語であるが、早くから言語共同体が成立し、語彙の使用が生産的な造語規則と言語共同体内の慣用との両方から規定される性質をもっているという点、および語構成の透明性が極めて高いという点において、言語の構造として特徴的であり、コロケーションの在り方に関しても示唆するところが多い。

平成22年度に予定されている国立国語研究所による「現代日本語書き言葉コーパス」の構築は、日本語コーパス言語学の進展、とりわけコロケーション研究への応用にとって大いに期待されることである。しかしながら、理論的な基盤に対する理解が伴ってはいじめ、コーパスを十分に活用することが可能になることは明らかである。コロケーションに関しても辞書に関する実践的蓄積は多いとは言えず、理論的基盤もまだ確固としては存在しない。

2. 研究の目的

本研究においては、コロケーションに関連する、語彙項目間のさまざまな関係の在り方を、コーパスを用いてより明確化するための通言語的な方法論の基礎を考究することを目指す。本研究課題の初期においては、英語研究一般、特にコーパス語彙研究において、コロケーション類似形式がこれまで理論的また実際的にどのような形に扱われてきたかを整理することが目標となる。

本研究課題の第二の目標は、計画言語エスペラントの語彙体系にみられる図式性と慣習性を整理した上で、それらの拮抗がコロケーションにどのような形で反映されているかを探る点にある。さらに、個別言語におけるコロケーション研究を、それぞれの言語における語の性質に留意しながら、通言語的コロケーション研究の基礎付けを行うことを目的とする。

3. 研究の方法

(1) 語彙論、辞書学の理論、特に英語コーパス言語学の立場からの語彙論、辞書学の分野への応用について、文献調査および研究者との意見交換を進め、コロケーション類似現象の理論的解明および辞書における記述のケース・スタディ的な調査を行った。その際に、フレイジオロジーと呼ばれて近年注目されつつある現象については、ドイツ語やロシア語に関する研究もみられることから、これらの言語圏における研究の流れにも特に留意した。

(2) エスペラントの計画性と社会性の関係について、世界エスペラント協会、日本エスペラント学会図書館等における文献調査を行った。また、エスペラントの社会性を重視す

べきことを確認するためオーストリア国立図書館を訪問して日本では入手困難な文献を閲覧した。同図書館では、さらにエスペラントおよびドイツ語圏、東欧圏におけるコロケーション研究に関する文献の調査を行った。

また、既存のエスペラントのテキストデータベースの量的な拡大に努める一方で、そこからのコロケーション情報の抽出の手法および言語教育上で有意義な情報の整理について、コーパスツールの利用方法および検索結果からのコロケーション情報の抽出について多数のケーススタディを行った。また、オランダ、アムステルダム大学を訪問して、テキストデータベースの構築と利用法について資料を収集し、研究者と意見を交換し、そこから得られた知見を随時ケーススタディに反映させた。

これらの研究を進めながら、内外の研究者と意見を交換しつつ、エスペラントのコロケーション研究の理論化に取り組んだ。

(3) 日本語に関しては、既存の文献とデータを利用して、コロケーション類似概念の分類について、理論的研究および実践的なパイロット調査を進めた。特に英語の場合との違いを念頭におきつつ、文献によりコロケーション研究の情報収集を継続した。英語、日本語、エスペラントそれぞれの言語構造の違いに応じたコロケーションの位置づけについて、考察を深めた。

4. 研究成果

(1) 英語コーパス言語学に関しては、語彙論、辞書学の分野への応用についての文献研究およびその他の情報収集において、フレイジオロジー、コリゲーション、イディオムなどのコロケーション類似現象に特に注目したが、ドイツ語、ロシア語におけるフレイジオロジー研究の流れとの関連性が当初の予想以上に大きいことが確認できた。

研究の蓄積が大きいものの、派生語におけるコロケーションの継承の有無については、理論的に未決着の部分が多く、今後の研究が必要であることが分かった。

(2) 文献調査の結果、エスペラントの成立における語根レベルでの計画性を強調しがちな従来の固定観念は適切なものではなく、構文からテキストに及ぶレベルにおける社会的慣習性をさらに重視すべきとの見解を得るにいたった。

テキストデータベースからのコロケーション情報の抽出の手法および言語教育上で有意義な情報の整理について行ったケーススタディから、語義区分ごとのコロケーションの異同、また派生に際してのコロケーシ

ンの継承の有無について、個別語彙項目ごとに一定の知見を得ることができ、その成果は随時発表した。その教育上の応用についても一定の成果を得た。

語が基本的な語義から転義的な語義に意味が拡張されている場合、その意味の変化に応じてコロケーションのパターンも変わるが、これは動詞や形容詞において顕著に見ることができる。

例えば、動詞 danki 「感謝する」は、その意味からして実生活においてさまざまな場面で用いられるが、くだけた場面と堅苦しい場面との使用の違いは、副詞(句)とのコロケーションの違いと関係している。

また、この動詞を修飾する副詞と同じ語根から作る名詞 danko 「感謝」を修飾する形容詞には、ある程度の平行性が見られる。しかし、このような派生に際してのコロケーションの継承は、機械的に見られるわけではない。動詞 zorgi 「気遣う、心配する」を修飾する強意副詞は、*multe, tre, tro, bone* が多く、*atente, diligente* なども見られるが、名詞 *zorgo* を修飾する強意形容詞は *granda, grava, profunda, forta* などであり、平行関係が見られない。

さらに、形容詞 *zorga* は「細心の、入念な」の語義で *preparo, kontrolo, analizo, esploro, laboro* などの名詞との共起が認められる。一方、副詞 *zorge* は対象に対する働きかけを意味する *atenti, aŭskulti, rigardi* や知的活動や判断を意味する動詞 *distingi, ekzameni, esplori, elekti, eviti, konsideri, kontroli, legi, observi, studi, pripensi*、準備や仕上げを意味する動詞 *aranĝi, prepari, ellabori, fari*、身づくろいや清掃に関する動詞 *balai, brosi, kombi, lavi, razi* などとのコロケーションが目立つ。このように品詞の転成においてコロケーションが必ずしも形式的な平行関係を示さない事例はほかにも多く見られた。それぞれの語の意味および頻繁に使われる使用場面などの要因が関わっている。

エスペラントは形態論的な透明性が高いため、これらの現象を観察することが容易であるという事実は特筆に値する。他の言語においては、品詞の転成において別語根の語が現れることがあるため、コーパスツールの利用においても検索結果の読み取りにおいても、現象の観察が容易ではない。また、コロケーションのパターンに違いがみられたにしても、その要因を特定することが難しい。

(3) 日本語については、国立国語研究所による現代日本語書き言葉均衡コーパスが完成したことから、その標準的な検索システムである少納言および中納言を利用したコロケ

ーションの抽出方法について検討を行い、正規表現の一定程度の有用性が確認できた。DVD版については、公開が遅れたので、十分な検討をすることができなかったが、さまざまな条件でコロケーションを抽出するための検索ツールの整備が今後は不可欠である。

(4) これらの結果、コロケーション類似現象は統語的に密接な関係にある内容語同士の場合を典型としつつも、その周辺部においては、それぞれの言語における語の形態論的性質に依存して、多様な広がりを見せていて、コーパスからのコロケーションの抽出にあたっては、このことを十分に考慮すべきであるとの結論にいたった。また、統計的な調査に加えて、言語体系および運用上の位置づけと文化的な意味づけを考慮すべきであることも確認された。これらの観察はコーパスを言語教育および辞書編集に利用する際に於いて念頭に置かれるべきことである。

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 28 件)

1. 後藤 斉、「de vorto al vorto (28) kreski」『エスペラント』第 80 巻(2012)4 月号, pp.11-12. 査読なし。
2. 後藤 斉、「de vorto al vorto (27) ĝui」『エスペラント』第 80 巻(2012)3 月号, pp.9-10. 査読なし。
3. 後藤 斉、「de vorto al vorto (26) nepre」『エスペラント』第 80 巻(2012)2 月号, pp.9-10. 査読なし。
4. 後藤 斉、「de vorto al vorto (25) ŝuldi」『エスペラント』第 80 巻(2012)1 月号, pp.12-13. 査読なし。
5. 後藤 斉、「de vorto al vorto (24) morti」『エスペラント』第 79 巻(2011)11 月号, pp.11-12. 査読なし。
6. 後藤 斉、「de vorto al vorto (23) plori」『エスペラント』第 79 巻(2011)10 月号, pp.9-10. 査読なし。
7. 後藤 斉、「de vorto al vorto (22) peza」『エスペラント』第 79 巻(2011)7 月号, pp.23-24. 査読なし。
8. 後藤 斉、「de vorto al vorto (21) pluvo」『エスペラント』第 79 巻(2011)6 月号, pp.11-12. 査読なし。
9. 後藤 斉、「de vorto al vorto (20) okazi」『エスペラント』第 79 巻(2011)4 月号, pp.11-12. 査読なし。
10. 後藤 斉、「de vorto al vorto (19) kosti」『エスペラント』第 79 巻(2011)3 月号, pp.11-12. 査読なし。
11. 後藤 斉、「de vorto al vorto (18) forta」

『エスペラント』第79巻(2011)2月号, pp.11-12. 査読なし。

12. 後藤 斉、「de vorto al vorto (17) flui」『エスペラント』第79巻(2011)1月号, pp.15-16. 査読なし。

13. 後藤 斉、「de vorto al vorto (16) riĉa」『エスペラント』第78巻(2010)12月号, pp.12-13. 査読なし。

14. 後藤 斉、「de vorto al vorto (15) zorgi」『エスペラント』第78巻(2010)11月号, pp.12-13. 査読なし。

15. 後藤 斉、「de vorto al vorto (14) maniero」『エスペラント』第78巻(2010)10月号, pp.12-13. 査読なし。

16. 後藤 斉、「de vorto al vorto (13) tuj と baldaŭ」『エスペラント』第78巻(2010)7月号, pp.14-15. 査読なし。

17. 後藤 斉、「de vorto al vorto (12) profunda」『エスペラント』第78巻(2010)6月号, pp.12-13. 査読なし。

18. 後藤 斉、「de vorto al vorto (11) flanko」『エスペラント』第78巻(2010)5月号, pp.12-13. 査読なし。

19. 後藤 斉、「de vorto al vorto (10) peni」『エスペラント』第78巻(2010)4月号, pp.12-13. 査読なし。

20. 後藤 斉、「de vorto al vorto (9) bela」『エスペラント』第78巻(2010)3月号, pp.12-13. 査読なし。

21. 後藤 斉、「de vorto al vorto (8) danki」『エスペラント』第78巻(2010)2月号, pp.7-8. 査読なし。

22. 後藤 斉、「de vorto al vorto (7) klara」『エスペラント』第78巻(2010)1月号, pp.14-15. 査読なし。

23. 後藤 斉、「de vorto al vorto (6) vera」『エスペラント』第77巻(2009)12月号, pp.6-7. 査読なし。

24. 後藤 斉、「de vorto al vorto (5) densa」『エスペラント』第77巻(2009)11月号, pp.6-7. 査読なし。

25. 後藤 斉、「de vorto al vorto (4) multa」『エスペラント』第77巻(2009)10月号, pp.6-7. 査読なし。

26. 後藤 斉、「de vorto al vorto (3) varma」『エスペラント』第77巻(2009)8月号, pp.6-7. 査読なし。

27. 後藤 斉、「de vorto al vorto (2) adiaŭ と Ĝis revido」『エスペラント』第77巻(2009)7月号, pp.6-7. 査読なし。

28. 後藤 斉、「de vorto al vorto (1) afabla」『エスペラント』第77巻(2009)6月号, pp.8-9. 査読なし。

[学会発表] (計 1 件)

後藤 斉「コロケーションを考えるためのいくつかの視点」日本エスペラント学会 2010

年度研究発表会、2010年10月9日、長崎

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

名称:

発明者:

権利者:

種類:

番号:

出願年月日:

国内外の別:

○取得状況 (計 0 件)

名称:

発明者:

権利者:

種類:

番号:

取得年月日:

国内外の別:

[その他]

ホームページ等

<http://www.sal.tohoku.ac.jp/~gothit/got-hitj.html> からリンクしている。

6. 研究組織

(1) 研究代表者

後藤 斉 (GOTOO HITOSI)

東北大学・大学院文学研究科・教授

研究者番号: 90162156

(2) 研究分担者

なし ()

研究者番号:

(3) 連携研究者

なし ()

研究者番号: