

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 23 日現在

機関番号：13801

研究種目：若手研究(B)

研究期間：2012～2014

課題番号：24720252

研究課題名(和文) 英語学習者と日本語学習者におけるコロケーションの情報処理過程に関する比較研究

研究課題名(英文) Parallel Processing of Collocations: A Comparative Study between Native Speakers and Non-Native Speakers

研究代表者

松野 和子 (Matsuno, Kazuko)

静岡大学・大学教育センター・講師

研究者番号：80615790

交付決定額(研究期間全体)：(直接経費) 1,300,000円

研究成果の概要(和文)：本研究による実験の結果、母語の違いに依らず、母語話者は概して二重処理によってコロケーションを処理するが、学習者は異なる処理を並行して同時に行うことが困難であることが考察された。特に、母語と異なる統語形式となるコロケーションは処理が困難であることが示唆された。学習者を習熟度別に分析すると、中級学習者に比べ、上級学習者では全体処理が行われるコロケーションが多くなることが分かった。表現パターンの記憶形成は、言語に内在する属性というよりは、経験によって得られた暗示的知識や明示的学習が影響していると考えられた。一方、コロケーションに内在する属性は、どの処理方法を採用するかに関係していることが示唆された。

研究成果の概要(英文)：The purpose of the present study is to examine whether native speakers and non-native speakers access pre-constructed patterns and those of each single word in a parallel way when the meanings of collocations are retrieved. The current study discusses collocational processing not only holistically or analytically but also "holistically and analytically." In the experiments conducted on native speakers of English/Japanese and learners of English/Japanese, the following conclusions were obtained: (1) native speakers generally process collocations in parallel but some kinds of collocations are solely processed; (2) learners solely process collocations and utilize less prefabricated patterns. It appears to be challenging for learners to run several processes simultaneously. This study suggests that experiences such as implicit knowledge and explicit learning influence memory formation of prefabricated patterns and expression characteristics impact what processing methods are adopted.

研究分野：第二言語習得

 キーワード：コロケーション 言語情報処理過程 コーパス 反応時間 プライミング 言語転移 第二言語習得 m
ulti-word units

1. 研究開始当初の背景

コロケーションは個々の単語の意味を組み合わせて全体の意味を理解できるが使用される語が慣習的に決まっている表現である。コロケーションを扱う際は、既に決まっている表現パターンに基づいて使うべき語が自動的に決まるため、言語を負担なく流暢に扱うことができる(Wray, 2002)。そのため、母語話者だけでなく外国語学習者にとっても、コロケーションは言語を扱う上で、重要な役割を果たす。

表現パターンが記憶されている語のまとまりを用いて理解や産出を行うことは言語情報処理過程として稀ではない(Erman & Warren, 2000)。意味の非分解性や統語規則からの逸脱など語のまとまりが持つ性質のために前もって記憶された表現パターンを用いて処理が行われる場合と、言語使用の経験によって記憶されるようになった表現パターンを用いて処理が行われる場合があるが、イディオム、句動詞、複合語、高頻度で用いられる語のまとまりは、記憶された表現パターンを用いて処理されることが明らかになってきている。

学習者のコロケーションの情報処理過程を明らかにする試みが近年行われ始めており、Wolter and Gyllstad (2011)では、学習者がコロケーションを理解する際は、コロケーション表現全体を母語に置き換える処理をしていることが分かった。しかしながら、コロケーションの情報処理過程について明らかになっていない点は依然として多い。例えば、学習者におけるコロケーションの情報処理過程の研究では、母語に置き換える処理をしているか等、情報処理過程の一部が明らかになっているが、コロケーション表現を聞いてからその意味を理解するまでに行われる情報処理過程の全体像は明らかになっていない。

2. 研究の目的

本研究の目的は、母語話者の情報処理過程と比較しつつ、英語学習者と日本語学習者におけるコロケーションの情報処理過程を明らかにすることである。その際、処理過程に関する全体像のモデルに基づき、学習者の情報処理過程を考察する。

3. 研究の方法

本研究では、第一に、コロケーションの定義について考察した。第二に、(1)語彙の情報処理過程、(2)文の情報処理過程、(3)イディオムの情報処理過程と統語処理に関する先行研究に基づき、コロケーションの情報処理過程のモデルを構築した。第三に、イディオムの情報処理過程を分析する手法を応用して、英語学習者、英語母語話者、日本語学習者、日本語母語話者を対象に以下の2

つの実験を行った。

実験1では、前もって記憶された表現パターンを用いて、コロケーションが処理されるかを検証するため、コロケーションの反応時間と自由連結表現の反応時間を比較した。その際、コロケーションと自由連結表現は同様の名詞によって構成されるように設定し、コロケーション内の動詞と自由連結表現内の動詞は同じ速度で処理されるように統制された。(カウンターバランスをとっているため、実験参加者は、同様の実験内でコロケーションとそれに対応する自由連結表現を目にすることはない。)

実験2では、コロケーションを処理する際、一単語ずつの意味を引き出しているかを検証するため、コロケーションが自由連結表現のいずれかを呈示した後に、コロケーションを構成している一単語と連想関係にある単語を呈示し、呈示された単語が同一の語(例: *wind*)であるのにもかかわらず先に呈示された表現(例: *blow my nose* または *hurt my nose*)によって反応時間に差が生じるかを分析した。(カウンターバランスをとっているため、実験参加者は、同様の実験内で、ターゲットとして同じ単語を繰り返し目にすることはなく、プライムとしてコロケーションとそれに対応する自由連結表現の両方を目にすることはない。)

分析の際は、表現に対する親密度や学習経験等を調査するアンケート結果とコロケーションに内在する言語表現の属性に基づき、同様の処理過程を持つコロケーションに共通する特徴や記憶の形成に影響する特徴を考察することを試みた。

4. 研究成果

第一に、本研究では、コロケーションを定義した。コロケーションを定義するため、イディオムとコロケーションを区別する基準を考察した結果、本研究では、2つの判別基準によってコロケーションとイディオムを定義することとした(Benson, 1985; Melcuk & Zholkovsky, 1988; Kjellmer, 1991; Cermak, 1994; Fraser, 1994; Lewis, 2000; Shei & Pain, 2000; Boers & Demecheleer, 2001; Abel, 2003)。

1 つめの基準は、慣習的に共に用いられる語が決まっているかどうかである(Gitsaki, 1999)。

2 つめの基準は、表現を構成している個々の語から表現の意味を理解できるかどうかである(Bolinger, 1976)。本研究では、コロケーションとイディオム両者とも、慣習的に共に用いられる語が決まっている表現と定義した。例えば、「死ぬ」という意味を表現する際に、*kick* と共に使用される語は *the bucket* と決まっている。そのため、*kick the bucket* は「慣習的に共に用いられる語が決まっている」という1つめの基準を満たすと考えられる。また、「決定する」という意味を表現する際、*decision* と共に *make* を使用すること

が決まっているため、*decision* に対して *make* を用いることは「慣習的に共に用いられる語が決まっている」という1つめの基準を満たすと考える。

本研究では、語の組み合わせを構成する個々の語から、表現の意味を予想することができるか否かによってコロケーションとイディオムを区別する。Benson (1985)によれば、個々の語の意味を組み合わせた表現全体の意味で比喩の意味を持ち、表現全体で1つの意味を有する表現がイディオムである。例えば、*kick the bucket* は、*kick, the, bucket* という1つ1つの単語の意味を組み合わせた表現全体の意味(「バケツを蹴る」)が比喩的な意味を持ち、表現全体で「死ぬ」という意味を持つ。表現を知らないと、*kick the bucket* の「死ぬ」という意味を理解できないため、*kick the bucket* はイディオムに分類される。一方、個々の単語から表現全体の意味が推測できるものがコロケーションである。例えば、*decision* に対して *make* が使用された場合、全体として比喩的な意味を有さず、*make*(～を作る)、*decision*(決定)という個々の単語の意味から、表現の意味である「決定する」を理解できるため、コロケーションと分類される。

コロケーションの慣習さを判定する操作的な基準は、コロケーションのコアな定義である「語と語が共に出現していること」(Sinclair, 1991)を基盤にして、主に以下の3つの基準のいずれか(または複数)が採用される。

- (a) 共起頻度
- (b) 結びつきの強さ
- (c) 類義語との代用性

また、以上の3つの基準に、以下の2点加わることがある。

- (d) 統語的つながり
- (e) 連続性

Kjellmer (1991)はコロケーションを「統語的つながりを持った、頻繁に用いられる語の連続である」と定義している ("Collocations are defined as recurring sequences that have grammatical structure.") (p. 116)。この Kjellmer (1991)の定義では、「(a) 共起頻度」と「(d) 統語的つながり」、「(e) 連続性」の3つの操作的な基準が用いられていることが分かる。例えば、*Although he, try to, green idea* という2語の組み合わせでは、すべてが連続した語の組み合わせであるため、*Although he, try to, green idea* は「(e) 連続性」の基準を満たしている。しかし、*Although he* は「(e) 統語的つながり」を持たないため、コロケーションとはみなされない。また、*green idea* は「(a) 共起頻度」の高い語の組み合わせではないため、コロケーションとみ

なされない。そのため、Kjellmer(1991)の定義では、*Although he, try to, green idea* のうち、*try to* のみがコロケーションとなる。

第二に、本研究では、コロケーションの情報処理過程モデルを構築した。コロケーションの情報処理過程モデルを構築する第一の手順として、本研究は、言語の情報処理過程に関する先行研究をまとめた。Townsend and Bever (2001)によれば、言語処理は、慣習と計算から成る。ここで、計算(computations)とは規則に基づく(rule-based)処理のことを示す(Townsend & Bever, 2001, p. 400)。文を処理する際、「計算」に基づき、一語ずつ単語を処理し計算することによって、予期しない語の組み合わせを理解することができ、自由に創造的な産出を行うことが可能となる(Sinclair, 1987, 1991)。「計算」に基づく処理に比べ、慣れた語や言語表現、文を処理する「慣習」に基づく処理では、言語情報処理のすべての過程が行われるわけではなく、処理が短縮されることが分かってきている。「慣習」に基づく言語処理として、前もって記憶されている表現パターン(memorized sequences; prefabricated expressions; expressions stored as a unit; prefabricated patterns)を用いる処理では、すでに出来上がっているパターン全体を1つのまとまりとして引き出して文の産出や理解を行う(Pawley & Syder, 1983)。そのため、創造的な文章を初めから作り上げたり理解するよりも、流暢な産出やすばやい理解を可能とする(Corder, 1973; Pawley & Syder, 1983; Kuiper & Lin, 1989; Kuiper, 2000; Singleton, 2000)。言い換えれば、一語ずつ単語を処理する場合に比べ、前もって記憶されている表現パターンを用いた処理では、情報処理を速く行うことができる(Van Lancker, 1975; Alternberg & Eeg-Olofsson, 1990)。

一方、語彙を理解する言語情報処理では、自動的に単語が認識され、記憶してある単語へのアクセスが行われる。単語へのアクセスが行われた際には、その単語の意味情報、統語情報(例: 統語的範疇)、音韻情報(正書法情報)、形態素情報が引き出される(Levelt 1989, 1993; Jiang 2000)。

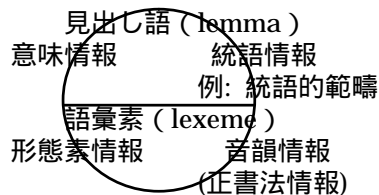


図1 単語へアクセスした際に引き出される情報

これらの単語の情報が引き出された後に、統語分析に基づき、表現を構成する単語の意味が計算され、表現や文の意味が引き出される。コロケーションの情報処理過程モデルを

構築する第二の手順として、イディオムにおける情報処理過程モデルの先行研究をまとめると、仮説 1「記憶されたパターンのみを用いて表現の意味を引き出す処理」(Direct Access Hypothesis: 直接アクセス仮説)、仮説 2「一語一語の意味を組み合わせて表現の意味を引き出す処理」(Idiom Decomposition Hypothesis: 分解処理仮説)、仮説 3「字義どおりの意味の後にイディオムの意味を引き出す処理」(Idiom List Hypothesis: イディオム表象仮説)、仮説 4「字義どおりの意味とイディオムの意味の 2 つの処理への着手時期が同時である二重処理」(Lexical Representation Hypothesis: 語彙的表象仮説)、仮説 5「ある語の並びがイディオムと認識された後に、表現パターンへのアクセスを開始する二重処理」(the Configuration Model: 形態認識仮説)に大きく分けられることが分かった。

イディオムの情報処理過程として、ある語の並びがイディオムと認識された後に、記憶された表現パターンへアクセスを開始し、記憶された表現パターンからも表現全体の意味を引き出しながら、並行して、一語ずつの単語の意味を処理して表現の字義的な意味を理解することが明らかになっている (Borrow & Bell, 1973; Swinney & Cutler, 1979; Estill & Kemper, 1982; Peterson & Burgess, 1993; Cacciari & Tabossi, 1988; Cacciari & Glucksberg, 1991; Titone & Connine, 1994)。また、イディオムの持つ属性(例: 意味の分解性、表現の出現頻度)や呈示のされ方(例: 文脈の有無、音調)によって情報処理過程が異なり、表現パターンだけを用いて意味が理解される場合や、一語ずつの意味を組み合わせるだけで表現の意味が理解される場合があると考えられている (Gibbs, 1986; Gibbs, Nayak & Cutting, 1989; Connine, Blasko, Brandt & Kaplan, 1992; Cronk, Lima & Schweigert, 1993; Glucksberg, Brown & McGlone, 1993; Abel, 2003)。加えて、イディオムと統語処理について、イディオムがひとまとまりで記憶されるかどうかに関しては統語の凍結度に関係なく、凍結度が低いイディオムもひとまとまりとして記憶されているが、凍結度によって処理のしやすさが異なることが明らかになっている (Swinney & Cutler, 1979; Gibbs & Gonzales, 1985)。また、統語分析が行われなくてもイディオムの意味を引き出すことができるが、通常は自動的に統語の処理が行われることが分かっている (Van Lancker & Kempler, 1987; Peterson & Burgess, 1993; Peterson, Burgess, Dell & Eberhard, 2001; Gagne & Spalding, 2004; Oliveri, et al. 2004)。

上記の先行研究に基づき、本研究では、コロケーションの処理過程モデルを構築し、以下の 3 つの情報処理過程の仮説のうちどの仮説に基づき、コロケーションの意味が理解されているかを検証した。

- (1) 前もって記憶された表現パターンのみを用いて、表現全体の意味を引き出す (全体処理モデル)
- (2) 一語ずつ単語の意味を処理して、表現全体の意味を引き出す (分解処理モデル)
- (3) ある語の並びが記憶された表現パターンであると認識された後に、表現パターンへのアクセスを開始して、(1)と(2)の両方の処理を行う (二重処理モデル)

図 2 は、コロケーションの情報処理過程における全体処理モデルを表したものである。

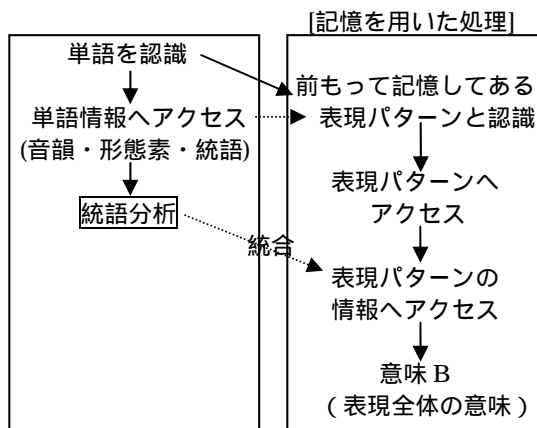


図 2 全体処理モデル

全体処理モデルでは、表現を構成する単語の認識が行われた後、それぞれの単語の情報へアクセスが行われる。単語の情報として、その単語の音韻情報、形態素情報、統語情報が含まれており、これらの情報が引き出される。この際に、品詞、語形変化、時制(の一部)も分析される。その後、表現の統語分析が行われる。このように、語を一語一語認識する一方で、並行して、表現を「前もって記憶された表現パターン」として認識する。その後、その表現パターンに関する情報にアクセスし、表現パターンが有する意味を引き出す。引き出された意味と統語分析が統合され、表現全体の意味が理解される。*answer*と*phone*が共起した表現を例とすると、*answer*と*phone*という単語の認識がそれぞれ行われた後、脳内に記憶されている各単語に関する情報へアクセスが行われ、それぞれの単語(*answer*・*phone*)の音韻情報、形態素情報、統語情報が引き出される。この際に、*answer*と*phone*の品詞、語形変化、時制が分析され、表現の統語分析が行われる。*answer*と*phone*という単語の認識がそれぞれ行われる一方で、並行して、*answer*と*phone*(*answer* + *phone*)がひとまとまりで認識され、*answer* + *phone*が脳内に記憶されている表現パターンであると認識される。その後、脳内に記憶されている *answer* + *phone* の情報にアクセスし、*answer* + *phone* が有している「電話に出る」という意味が引き出され、統語分析と

表現の意味(「電話に出る」)を統合して表現全体の意味が理解される。

図3は、コロケーションの情報処理過程における分解処理モデルである。

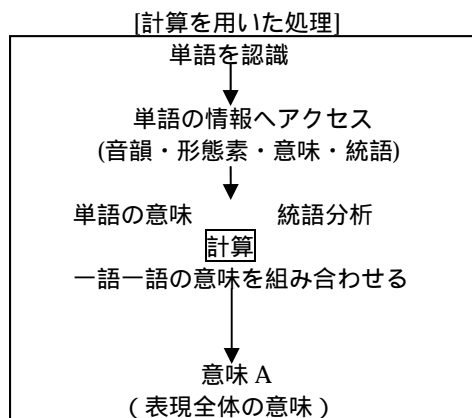


図3 分解処理モデル

分解処理モデルでは、表現を構成する単語の認識が行われた後、それぞれの単語の情報へアクセスが行われる。単語の情報では、単語の音韻情報、形態素情報、意味情報、統語情報が含まれており、これらの情報が引き出される。この際に、品詞、語形変化、時制(の一部)も分析される。その後、それぞれの語の意味と統語分析に基づく計算が行われ、表現全体の意味が理解される。この情報処理過程は、自由連結表現の意味理解と同様のモデルである。*answer* と *phone* が共起した表現を例とすると、*answer* と *phone* という単語の認識がそれぞれ行われた後、脳内に記憶されている各単語の情報へアクセスが行われ、それぞれの単語の音韻情報、形態素情報、意味情報、統語情報が引き出される。この際に、*answer* と *phone* の品詞、語形変化、時制も分析される。その後、それぞれの単語の意味(引き出された *answer* の意味・*phone* の意味)と統語分析に基づく計算が行われ、表現全体の意味が理解される。

図4はコロケーションの情報処理過程における二重処理モデルである。

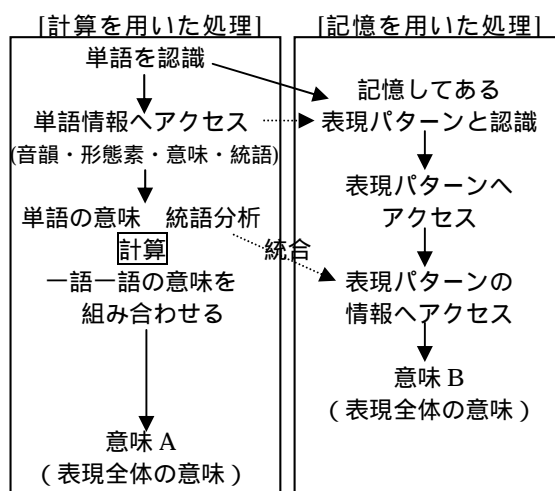


図4 二重処理モデル

二重処理モデルでは、表現を構成する単語の認識が行われた後、それぞれの単語へのアクセスが行われ、単語の情報へのアクセスが行われる。単語の情報では、各単語の音韻情報、形態素情報、意味情報、統語情報が含まれており、これらの情報が引き出される。この際に、品詞、語形変化、時制(の一部)も分析される。その後、それぞれの語の意味と統語分析に基づいて計算が行われ、表現全体の意味が理解される。語を一語ずつ認識する一方で、並行して、表現を「前もって記憶された表現パターン」であると認識する。その後、表現パターンに関する情報にアクセスし、表現パターンが有する意味を引き出す。表現パターンとして引き出された意味と統語分析が統合され、表現全体の意味が理解される。*answer* と *phone* が共起した表現を例とすると、*answer* と *phone* という単語の認識がそれぞれ行われた後、脳内に記憶されている各単語の情報へアクセスが行われ、それぞれの単語の音韻情報、形態素情報、意味情報、統語情報が引き出される。この際に、*answer* と *phone* の品詞、語形変化、時制も分析される。その後、それぞれの単語の意味(*answer* の意味・*phone* の意味)と統語分析に基づく計算が行われ、表現全体の意味が理解される。*answer* と *phone* という単語の認識が行われる一方で、*answer* + *phone* がひとまとまりで認識され、*answer* + *phone* が脳内に記憶された表現パターンであることが認識される。その後、脳内に記憶された *answer* + *phone* の情報にアクセスし、*answer* + *phone* が有している表現の意味(「電話に出る」)が引き出される。その後、統語分析と表現の意味(「電話に出る」)を統合して、表現全体の意味が理解される。二重処理モデルでは、計算による処理と記憶による処理の2つのルートによって、表現全体の意味が引き出されることとなる。

本研究では、以上の「全体処理モデル」(図2)、「分解処理モデル」(図3)、「二重処理モデル」(図4)のうち、どのモデルに基づいて、

英語母語話者・日本語母語話者と英語学習者・日本語学習者はコロケーションの意味を理解しているかを明らかにすることを試みた。実験によって得られたデータを分析した結果(実験手法については本稿の「3. 研究の方法」を参照)、母語の違いに依らず、同様の傾向がみられ、母語話者では概して二重処理によってコロケーションを処理するが、学習者は異なる処理を並行して同時に行うことが困難であることが考察された。特に、母語と異なる統語形式となるコロケーションを処理する場合、処理が困難であることが示唆された。学習者を習熟度別に分析すると、中級学習者では分解処理が行われ、一単語ずつの意味を組み合わせてコロケーションの意味を理解することが多かった。一方、上級学習者では、全体処理が行われるようになり、表現パターンを用いて意味を理解するコロケーションが多くなることが分かった。これは、中級学習者ではひとまとまりとして前もって記憶しておいた表現パターンが少なく、上級学習者ではひとまとまりとして前もって記憶している表現パターンが多くなるためであると考察された。表現パターンの記憶形成は、言語に内在する属性というよりは、経験によって得られた暗示的知識や明示的学習が影響していると考えられた。一方、コロケーションに内在する表現の属性は、どの処理方法を採用するかに影響を与えていると考察された。

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2件)

松野和子 (2015).「コロケーションにおける同義語の代用性に関する一考察」『静岡大学教育研究』第11号, 185-196. (査読有)

松野和子 (2014).「コロケーションの定義に関する再考」『静岡大学教育研究』第10号, 67-81. (査読有)

[学会発表](計1件)

Kazuko Matsuno 2014年9月4日.
Parallel processing of collocations: Do native speakers and second language learners access prefabricated patterns and that of each word simultaneously?
EURO-SLA 24,
The University of York (イギリス)
(査読有)

[その他]

ホームページ等

<https://tdb.shizuoka.ac.jp/RDB/public/Default2.aspx?id=10809&l=0>

6. 研究組織

(1) 研究代表者

松野 和子 (MATSUNO, Kazuko)

静岡大学・大学教育センター・講師

研究者番号：80615790